



TRABALHO DE GRADUAÇÃO

**MODELAGEM E SIMULAÇÃO DE ALGORITMO  
DE RASTREIO BASEADO EM FILTRO DE KALMAN  
PARA ESTIMAÇÃO DE CANAL EM  
SISTEMAS MIMO DE COMUNICAÇÃO**

Lúcio Sidney de Oliveira Souza

Brasília, Dezembro de 2017

**UNIVERSIDADE DE BRASÍLIA**

**FACULDADE DE TECNOLOGIA**

UNIVERSIDADE DE BRASÍLIA  
Faculdade de Tecnologia

TRABALHO DE GRADUAÇÃO

MODELAGEM E SIMULAÇÃO DE ALGORITMO  
DE RASTREIO BASEADO EM FILTRO DE KALMAN  
PARA ESTIMAÇÃO DE CANAL EM  
SISTEMAS MIMO DE COMUNICAÇÃO

Lúcio Sidney de Oliveira Souza

*Relatório submetido ao Departamento de Engenharia  
Elétrica como requisito parcial para obtenção  
do grau de Engenheiro Eletricista*

Banca Examinadora

Prof. João Paulo Leite , ENE/UnB \_\_\_\_\_  
*Orientador*

Prof. Leonardo R. A. X. de Menezes, ENE/UnB \_\_\_\_\_  
*Examinador interno*

Prof. Luís Afonso Bermúdez, ENE/UnB \_\_\_\_\_  
*Examinador interno*

## **Dedicatória**

*A minha família que muito me apoiou em todos os momentos.*

*Lúcio Sidney de Oliveira Souza*

## Agradecimientos

*Agradeço a Deus pela vida e pela saúde.*

*Agradeço aos meus pais por todo carinho e atenção que me deram. Obrigado, por terem investido tanta energia e esforço em minha educação para que fosse sempre um melhor ser humano. E, principalmente, por terem acreditado em mim e sempre terem feito grandes investimentos em meu ensino, não há presente maior a ser dado a uma pessoa do que um ensino de qualidade. Agradeço ao meu irmão por estar presente na minha vida e saber que sempre terei alguém com quem contar. A família é a base de tudo! Sem vocês não teria chegado até aqui.*

*Obrigado aos funcionários do Departamento de Engenharia Elétrica.*

*Obrigado a parcela de excelentes professores com quem me deparei no curso e que são verdadeiros exemplos de pessoas que amam o que fazem.*

*Agradeço especialmente ao professor Anésio de Leles Ferreira Filho por suas aulas, pelo espaço no projeto, por seus conselhos, pelas conversas, pelas reuniões e por toda a confiança. Sou muito grato, pois, muito aprendi no Laboratório de Qualidade de Energia Elétrica.*

*Segundo Galileu Galilei não se pode ensinar tudo a alguém. Pode-se, apenas, ajudá-lo a encontrar por si mesmo. E, foi isso o que encontrei nesse trabalho com o professor João Paulo Leite como orientador. Muito obrigado por todas as horas de reunião, por todas as dúvidas sanadas, pelas conversas, pelas orientações e, principalmente, pela paciência. Agradeço a toda cobrança feita nas disciplinas e no trabalho de conclusão de curso, pois, com certeza, muito me fizeram crescer e ser uma pessoa melhor.*

*Obrigado as minhas amigas Anna Eloyr, Marina Pinheiro e Renata Oliveira, obrigado por sempre me ouvirem e aconselharem. Passamos muito tempo juntos, momentos decisivos em nossas vidas e, principalmente, bons momentos. Vocês são pessoas incríveis que agradeço sempre por ter vocês. Que a nossa amizade dure para sempre.*

*Agradeço também as minhas amigas Caroline Cunha e Débora Ferreira por todos os momentos em que me apoiaram, aconselharam e ajudaram. Vocês foram importantes nessa etapa final para que esse trabalho fosse possível.*

*Obrigado a todos os meus colegas e amigos de curso pelo apoio, conversas, alegrias e angústias compartilhadas. Passamos por diversos momentos difíceis, situação inimagináveis! Obrigado especialmente aos que entraram comigo nessa longa jornada: Antonio Alves, Bruno Floriano, Guilherme Torres, Jéssica Marie, Jéssica Holanda, Kássia Sayonnara, Jacqueline Cristina, Lukas Bezerra e Victor Brasil. Com toda certeza vocês fizeram parte da minha história e nunca serão esquecidos. São não apenas pessoas exemplares como futuros engenheiros e engenheiras, mas também como seres humanos. Obrigado novamente, vocês deram sentido a engenharia quando não havia!*

*Lúcio Sidney de Oliveira Souza*

---

## RESUMO

Nas últimas décadas, houve um aumento do uso de comunicações móveis, como consequência novas demandas foram criadas como maior tráfego de dados, maior confiabilidade, menor custo e menor latência. Esse fato faz necessário novas ferramentas que satisfaçam essas especificações, já que essas não podem ser atendidas simplesmente pelo aumento a largura de banda, que é um recurso escasso, ou da potência de transmissão, que levaria a um aumento da interferência entre sistemas, eleva custos de implementação e operação do sistema. A comunicação 5G por meio do *Massive multiple-input multiple-output* (MIMO Massivo) mostra-se, então, como uma das soluções a essa questão.

Para que os benefícios da utilização de um grande número de antenas sejam ampliados em sistema MIMO é necessário o conhecimento do estado do canal, *channel state information* (CSI). Geralmente, melhores desempenhos são possíveis quando transmissores e receptores têm informações sobre o canal de comunicação. Este trabalho trata do caso em que apenas o transmissor tem informação sobre o canal de comunicação, ou seja, um sistema em malha aberta operando em enlace reverso.

Neste contexto, esse trabalho faz a comparação entre dois algoritmos para o rastreamento de canais de comunicação que apresentam desvanecimento plano, o algoritmo *least mean square* (LMS) e o filtro de Kalman. Para tanto, foram feitas simulações para o rastreamento de um canal baseado no modelo de canal de Jakes, em que se notou um melhor desempenho do filtro de Kalman frente ao algoritmo LMS.

**Palavras-chave::** MIMO massivo, canal de desvanecimento plano, filtro de Kalman, algoritmo LMS, comunicações móveis.

---

## ABSTRACT

In the last decades, there was growth in the utilisation of mobile communications, as a result new demands has been created such as growing traffic data, greater reliability, lower cost and a lower latency. This fact creates the necessity of new technologies that meet this especification, once those can not be satisfied just with a increase in the bandwidth, that is scarce resource, or in power transmission, that would result in a bigger interference between systems, higer cost of implementation and operation of the systems. The communication with 5G technology by Massive multiple-input multiple-output is presented as one of solutions proposed to this problem.

The knowledge of channels state informations are required so that the benefits of the utilisation of a large number of antennas are fully exploited. Generally, better performances are reached when transmitter and receptor have CSI. This work presents the case in which just the transmissor have information about the communication channel, that could be understood as an open loop system that operates in uplink.

In this context, this work does the comparation between two algorithms for tracking a channel which is classified as flat fading, the least mean square algorithm (LMS) and the Kalman filter. For this purpose, simulations were made for tracking a channel based on jakes model, in which it was noted a better performance of the Kalman filter when compared to the LMS algorithm.

**Index terms:** Massive MIMO, flat fading channel, Kalman filter, LMS algorithm, mobile communication.

# SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO .....</b>	<b>1</b>
1.1	CONTEXTUALIZAÇÃO .....	1
1.2	DEFINIÇÃO DO PROBLEMA .....	2
1.3	OBJETIVOS E ESTRUTURA DO TRABALHO .....	3
<b>2</b>	<b>SISTEMAS MIMO .....</b>	<b>5</b>
2.1	INTRODUÇÃO .....	5
2.2	CLASSIFICAÇÃO DOS SISTEMAS MIMO .....	5
2.3	DIVERSIDADE ESPACIAL E MULTIPLEXAÇÃO ESPACIAL .....	6
2.4	FENÔMENOS EM CANAIS MULTIPERCURSO .....	7
2.5	DESVANECIMENTO EM GRANDE ESCALA .....	8
2.6	DESVANECIMENTO EM PEQUENA ESCALA .....	9
2.7	PARÂMETROS RELACIONADOS AO ATRASO .....	10
2.7.1	ESPALHAMENTO DE ATRASO ( $T_m$ ) .....	11
2.7.2	BANDA DE COERÊNCIA ( $B_c$ ) .....	12
2.8	PARÂMETROS RELACIONADOS AO TEMPO .....	13
2.8.1	DESVIO DOPPLER ( $B_d$ ) .....	13
2.8.2	TEMPO DE COERÊNCIA ( $T_c$ ) .....	13
2.8.3	CLASSIFICAÇÃO DOS CANAIS MULTIPERCURSO .....	14
2.9	MODELO DO SISTEMA .....	15
2.10	MODELO DE SISTEMA COM CORRELAÇÃO .....	20
2.11	CONCLUSÕES PARCIAIS .....	20
<b>3</b>	<b>MODELO DE CANAL .....</b>	<b>22</b>
3.1	INTRODUÇÃO .....	22
3.2	DESVIO DOPPLER E MODELO DE CANAL DE CLARK .....	22
3.3	MODELO DE CANAL DE JAKES .....	24
3.4	DESENVOLVIMENTO MATEMÁTICO .....	25
3.5	CONSTRUÇÃO DO MODELO .....	27
3.6	MODELAGEM DO CANAL POR MEIO DA CONFORMAÇÃO ESPECTRAL .....	29
3.7	RESULTADOS PARCIAIS E CONCLUSÕES .....	30
<b>4</b>	<b>ALGORITMO LMS PARA RASTREIO DE CANAL .....</b>	<b>35</b>



4.1	INTRODUÇÃO .....	35
4.2	PROBLEMA PARA REDUÇÃO DO ERRO QUADRÁTICO MÉDIO.....	35
4.3	MÉTODO DO GRADIENTE DESCENDENTE.....	36
4.4	ALGORITMO LMS .....	39
4.5	CURVA DE APRENDIZADO .....	41
4.6	RESULTADOS.....	42
4.7	CONCLUSÕES PARCIAIS.....	48
<b>5</b>	<b>FILTRO DE KALMAN PARA RASTREIO DE CANAL .....</b>	<b>49</b>
5.1	INTRODUÇÃO .....	49
5.2	DEMONSTRAÇÃO INTUITIVA DO FILTRO DE KALMAN .....	49
5.3	DESCRIÇÃO MATEMÁTICA DO FILTRO DE KALMAN.....	54
5.4	FILTRO DE KALMAN APLICADO A ESTIMAÇÃO DE CANAIS.....	55
5.5	RESULTADOS.....	57
5.6	CONCLUSÕES PARCIAIS.....	59
<b>6</b>	<b>CONCLUSÃO .....</b>	<b>63</b>
6.1	COMENTÁRIOS SOBRE OS RESULTADOS OBTIDOS.....	64
6.2	SUGESTÕES PARA TRABALHOS FUTUROS.....	64
	<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>66</b>
	<b>ANEXOS.....</b>	<b>68</b>
<b>I</b>	<b>DEMONSTRAÇÃO DAS EQUAÇÕES DO FILTRO DE KALMAN.....</b>	<b>69</b>
I.1	ESTÁGIO DE PREDIÇÃO .....	69
<b>II</b>	<b>CONVERGÊNCIA DO ALGORITMO LMS .....</b>	<b>71</b>
II.1	CONVERGÊNCIA NA MÉDIA .....	71
II.2	CONVERGÊNCIA NA MÉDIA QUADRÁTICA .....	73

# LISTA DE FIGURAS

1.1	Representação de um sistema em malha fechada [1].....	2
2.1	Possíveis configurações de antenas e nomenclatura atribuídas ao sistema. Adaptado de [2]. .....	6
2.2	Fenômenos responsáveis por criar multi-percursos. Adaptado de [2]. .....	8
2.3	Representação do desvanecimento de grande escala e de pequena escala. Adaptado de [2]. .....	8
2.4	Representação da função resposta ao impulso de um sistema variante no tempo no domínio do tempo e do atraso. Adaptado de [2]. .....	10
2.5	Espectro de atraso de potência. Adaptado de [2]. .....	12
2.6	Par transformada de Fourier: atraso no tempo e banda de coerência. Adaptado de [2]. .....	13
2.7	Par transformada de Fourier: espalhamento Doppler e tempo de coerência. Adaptado de [2]. .....	14
2.8	Modelo de sistema MIMO. Adaptado de [3]. .....	15
2.9	Capacidade ótima para sistema MIMO $4 \times 4$ , SNR = 10 dB e 1000 realizações.....	18
2.10	Comparações entre as capacidades do canal para matriz de ganhos descrita em Eq. (2.42).....	19
3.1	Representação do desvio Doppler. Adaptado de [2]. .....	23
3.2	Representação do modelo de canal de Clark [2]. .....	23
3.3	Espectro de Potência Doppler para portadora de frequência igual a 2,4 GHz.....	24
3.4	Angulo de chegada da componente de um espalhador.....	26
3.5	Representação da construção do modelo de Jakes. Adaptado de [4]. .....	28
3.6	Descrição do processo necessário para a criação de uma sequência aleatória com uma dada correlação. Adaptado de [5]. .....	29
3.7	Densidade espectral de potência para amostras independentes. Adaptado de [5]. .....	29
3.8	Tipos de filtros comumente usados. Adaptado de [6]. .....	31
3.9	Resposta ao Impulso do canal e PSD.....	31
3.10	PSD para entrada como um ruído e realização temporal do canal. ....	32
3.11	Resposta ao Impulso do canal.....	33
3.12	Autocorrelação do ruído gaussiano complexo em que as componentes reais e imaginárias são independentes.....	33

3.13 Autocorrelação do ruído gaussiano complexo em que as componentes reais e imaginárias são independentes após o processo de filtragem.....	34
4.1 Representação espacial do filtro.....	36
4.2 Representação da superfície de erro para caso de um peso apenas [7] .....	38
4.3 Representação em fluxo de sinais do algoritmo LMS [7]. .....	41
4.4 Rastreamento do canal de comunicação para taxa de aprendizado = 0,1. ....	43
4.5 Rastreamento do canal de comunicação para taxa de aprendizado = 0,2. ....	44
4.6 Rastreamento do canal de comunicação para taxa de aprendizado = 0,25.....	44
4.7 Erro quadrático médio pela taxa de aprendizado do algoritmo LMS.....	45
4.8 Erro quadrático médio pela SNR do algoritmo LMS. Para 100 realizações de canal de canais. $N_t = 2$ .....	46
4.9 Erro quadrático médio pela SNR do algoritmo LMS. Para 100 realizações de canal de canais. $N_t = 200$ .....	46
4.10 Variação da ordem do filtro (p). ....	47
5.1 Representação do problema. Adaptado de [8]. ....	51
5.2 Informações iniciais sobre o sistema no tempo $t = 0$ [8]. ....	52
5.3 Informações iniciais sobre o sistema no tempo $t = 1$ [8]. ....	52
5.4 Determinação da posição do trem com base em medições ruidosas sobre a posição e previsões passadas. Adaptado de [8]. ....	53
5.5 Ciclo do filtro de Kalman.....	56
5.6 Diagrama do rastreamento do canal de comunicação com o uso filtro de Kalman. ....	58
5.7 Corte feito no rastreamento do canal de comunicação com o uso filtro de Kalman.....	59
5.8 Erro quadrático médio entre saída do filtro de Kalman, $y_e$ , e planta do processo, $y$ pela SNR.....	60
5.9 Erro quadrático médio entre saída do filtro de Kalman, $y_e$ e canal real, $u$ , pela SNR. ....	60
5.10 Erro quadrático médio pelo coeficiente de desvanecimento temporal.....	61
5.11 Erro quadrático médio entre saída do filtro de Kalman, $y_e$ , e canal real, $u$ , pela SNR. ....	61

# LISTA DE TABELAS

# LISTA DE SÍMBOLOS

## Símbolos Latinos

$B_c$	Banda de coerência	[Hz]
$B_d$	Desvio Doppler	[Hz]
$\mathbf{B}_t$	Matriz de controle de entrada	
$c$	Velocidade da luz	[m/s]
$d$	Distância	[m]
$\mathbf{F}$	Matriz de transição	
$f$	Frequência	[Hz]
$G_t$	Ganho de transmissão	
$G_r$	Ganho de recepção	
$h$	Ganho do canal de comunicação	
$J_0(.)$	Função de Bessel do primeiro tipo e ordem zero	
$L$	Perdas	
$M$	Desajuste	
$N$	Número de espalhadores	[Unidades]
$p$	Sensores	[Unidades]
$P$	Potência	[Watts]
$\mathbf{Q}$	Matriz de covariância do ruído do processo	[Watts]
$r$	Amplitude da envoltória   Sinal recebido	[Watts] Watts]
$\mathbf{R}$	Matriz de covariância	
$s$	Sinal transmitido	[Watts]
$S$	Espectro Doppler	[Hz]
$t$	Tempo	[s]
$T_c$	Tempo de coerência	[s]
$T_m$	Espalhamento de atraso	[s]
$\mathbf{v}_t$	Vetor ruído de medições	
$w_d$	Frequência Doppler	[rad/s]
$w_t$	Pesos do filtro LMS	[Watts]
$\mathbf{x}$	Vetor de estados	
$z$	Ruído	[Watts]
$\mathbf{z}_t$	Vetor de medições	

## Subescritos

$c$	Portadora
$r$	Recepção
$t$	Transmissão
$fs$	Espaço livre ( <i>free space</i> )
<i>Outras</i>	Outras perdas do modelo

## Símbolos Gregos

$\alpha$	Coeficiente de desvanecimento temporal   Ângulo de chegada
$\kappa_t$	Matriz de ruído do processo
$\Lambda_t$	Matriz de transição
$\sigma$	Parâmetro escalar da distribuição   Desvio padrão
$\sigma_t$	Valor quadrático médio do máximo excesso de atraso
$\phi$	Função de correlação
$\mu$	Média
$\nu$	Frequência

## Notações e Operadores Matemáticos

$\cdot$	Derivada temporal
$\langle \rangle$	Média temporal
$   $	Absoluto
$\  \ $	Norma de Frobenius
$E$	Média
$T$	Transposto conjugado

## Siglas

AR	Auto-regressiva
ARMA	Modelos auto-regressivos de médias móveis ( <i>autoregressive-moving-average</i> )
BER	Taxa de erro de <i>bit</i> ( <i>Bit error rate</i> )
CSI	Informação de Estado do Canal ( <i>Channel state information</i> )
FDP	Função densidade de probabilidade
FIR	Resposta finita ao impulso ( <i>Finite impulse response</i> )
IIR	Resposta infinita ao impulso ( <i>Infinite impulse response</i> )
LMS	Mínimos quadrados médios ( <i>Least mean squares</i> )
LOS	Visada direta ( <i>Line-of-sight</i> )
MIMO	Múltiplas entradas múltiplas saídas ( <i>Multiple input multiple output</i> )
MISO	Múltiplas entradas e uma saída ( <i>Multiple input sinle output</i> )
NLOS	Sem visada direta ( <i>Non-Line-of-sight</i> )
PSD	Densidade espectral de potência ( <i>Power spectral density</i> )
SIMO	Única entrada e múltiplas saídas ( <i>Sinle input multiple output</i> )
SNR	Relação sinal-ruído ( <i>Signal-to-noise-Ratio</i> )
TDD	Divisão por duplexação no tempo ( <i>Time division duplexing</i> )

# Capítulo 1

## Introdução

*Este capítulo apresenta a motivação desse trabalho tendo como base o cenário atual de comunicação. Por fim, é apresentada a organização do manuscrito.*

### 1.1 Contextualização

O grande desenvolvimento tecnológico permitiu, nas últimas décadas, um aumento do uso de comunicações móveis e uma maior integração dessas ao cotidiano. Dentro desse contexto novas demandas foram criadas como maior tráfego de dados, maior confiabilidade, menor custo e menor latência[9, 10]. Esse fato faz necessário novas ferramentas que satisfaçam essas especificações, já que essas não podem ser atendidas simplesmente pelo aumento a largura de banda, que é um recurso escasso, ou da potência de transmissão, que levaria a um aumento da interferência entre sistemas, eleva custos de implementação e operação do sistema [11, 12]. A comunicação 5G por meio do *Massive Multiple-Input Multiple-Output* (MIMO Massivo) mostra-se, então, como uma das soluções a essa questão.

Por *Multiple-Input Multiple-Output* (MIMO) denomina-se uma coleção de técnicas de processamento de sinal desenvolvidas para a melhor o desempenho de sistemas de comunicação. Por meio do uso de um maior número de antenas, é possível um aumento do número de graus de liberdade, que podem ser interpretados como um aumento do número de canais de comunicação. Esses podem ser usados ora para a transmissão de diversas cópias da informação que se deseja transmitir por meio da codificação espacial, diversidade espacial, ora para a transmissão de parcelas da informação que são combinadas no receptor, multiplexação espacial, [2].

Por MIMO massivo entende-se ainda o emprego de arranjos com centenas de antenas ou mais, cujo objetivo é a potencialização dos benefícios obtidos com o uso do MIMO convencional [13]. Em sistemas MIMO massivo, é possível a redução da interferência entre usuários que ocorre no MIMO multiusuário [14]. Além disso, é possível a utilização de amplificadores de potência lineares e baratos já que a potência de transmissão de cada antena pode ser bastante reduzida. Sistemas MIMO permitem o suporte a um grande número de usuários simultaneamente uma vez que se tem um aumento do taxa de transmissão com esse esquema [13]. No entanto, ressalta-se que muitas



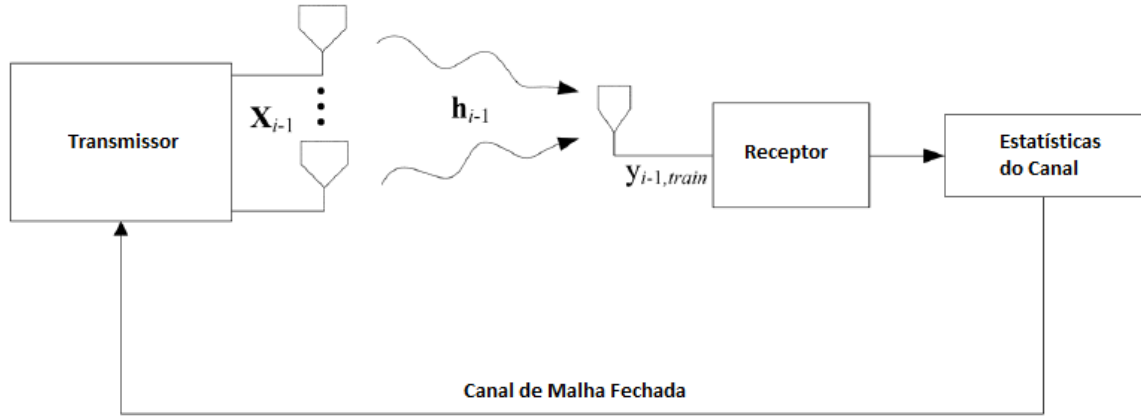


Figura 1.1: Representação de um sistema em malha fechada [1].

vezes esses benefícios não podem ser atingidos simultaneamente.

## 1.2 Definição do problema

Ambientes com grande espalhamento e utilização de técnicas MIMO possibilitam a criação de sistemas de comunicação móveis com maior capacidade, uma vez que permitem a multiplexação espacial da informação a ser transmitida. No entanto, para que as vantagens desse tipo de sistema possam ser utilizadas é necessário o conhecimento das informações do estado do canal - *channel state information (CSI)* - tanto pelo transmissor quanto pelo receptor. Portanto, o conhecimento das informações do canal são de extrema relevância para melhores desempenhos do sistema.

Ademais, as técnicas MIMO podem ser classificadas em duas categorias malha aberta e malha fechada [2]. Essa classificação tem como base o conhecimento do canal de comunicação pelo transmissor ou pelo receptor. A malha fechada ocorre quando o transmissor tem informações do canal de comunicação, isso é feito por meio dos sinais pilotos enviados pelo transmissor para o receptor que estima o canal de comunicação e os envia a estação base. A representação de um sistema em malha fechada é apresentado em Fig. 1.1. Este trabalho terá como foco o estudo em malha aberta.

Existem três grandes categorias em que são classificados os algoritmos de estimação de canal: cegos, semicegos e supervisionados [12, 15]. A primeira classe é caracterizada pela não transmissão de sinais de referência ou pilotos, com a qual é possível atingir uma maior eficiência espectral já que este todo será utilizado para transmissão de informação. O funcionamento desse algoritmo se baseia em propriedades e estatísticas do canal. A segunda classe se caracteriza pela transmissão de sinais pilotos e pelas estatísticas dos sinais transmitidos. O uso de processos cegos ou semicegos deve ser usada principalmente em sistemas de variação lenta em relação a taxa de transmissão de símbolos.

Por fim, a terceira forma de se estimar o canal - estimação supervisionada - é através da uti-

lização de sinais pilotos, que são um conjunto de símbolos transmitidos conhecidos previamente pela estação base e pelo usuário em que, com base nas informações recebidas, estima-se o canal de comunicação [13]. Nessa categoria de estimação considera-se que o canal permanecerá inalterado até a próximo envio de pilotos - caso por duplexação por divisão no tempo (*time division duplexing* - *TDD*). Esse sistema deve ser usado em sistemas de grande variação como é o caso da comunicação móvel em que, geralmente, há uma variação do canal causada pelo movimento relativo entre transmissor e receptor ou até mesmo mudanças no ambiente de propagação dos sinais.

As mudanças no canal devido ao movimento relativo podem ser interpretadas como sendo mudanças temporais. É preciso, portanto, a estimação do canal nesse ambiente de variações temporárias. Nesse contexto, o filtro de Kalman usado em processamento de sinais apresenta-se como uma das alternativas possíveis para a estimação dos canais. A sua base na minimização do erro quadrático médio associado ao comportamento recursivo eficiente permitem um bom desempenho desse filtro frente a outros métodos. Além disso, o modelamento de sistemas MIMO por meio de matrizes no espaço de estados permitem ainda uma melhor adequação do problema ao uso do filtro de Kalman.

### 1.3 Objetivos e Estrutura do trabalho

O objetivo desse trabalho é a análise da estimação de canais através do uso do filtro de Kalman frente ao algoritmo *Least Mean Squares* (LMS). Isso porque a correta estimação dos canais de comunicação é essencial para que os benefícios da utilização das técnicas MIMO sejam ampliados.

Os dois métodos foram escolhidos já que o filtro de Kalman é um algoritmo computacional que existe a mais de 50 anos, apesar disso, ainda possui grande utilização nos dias de hoje. Além de possuir o *status* de estimador ótimo de sistema lineares unidimensionais com estatísticas de erro gaussianas. Enquanto o algoritmo LMS remota a década de 60, é um algoritmo simples, não requer medidas da função de correlação do sistema nem a inversão da matriz de correlação.

Com base neste objetivo, tem-se a seguinte organização:

- Capítulo 2: É apresentado neste capítulo os principais conceitos em relação aos sistemas MIMO de comunicação que são fundamentais para o desenvolvimento dos capítulos subsequentes deste trabalho;
- Capítulo 3: Com a base teórica mostrada no Capítulo 2, há no capítulo 3 uma descrição do modelo de canal de comunicação que será usado para ser rastreado por meio do filtro de Kalman e do algoritmo LMS que serão descritos nos capítulos 4 e 5;
- Capítulo 4: Neste capítulo, há a apresentação de um exemplo ilustrativo para o desenvolvimento intuitivo do algoritmo LMS. Na sequência, é apresentado o desenvolvimento das equações usadas desse. Há ainda a presença dos resultados do rastreamento do canal usando esse método de rastreamento;

- Capítulo 5: Neste capítulo há a apresentação de um exemplo ilustrativo para o desenvolvimento intuitivo do filtro de Kalman. Em seguida, há a demonstração das equações presentes nesse algoritmo. Por fim, são mostrados os resultados referentes ao rastreamento de um canal de comunicação usando esse algoritmo e a comparação dos resultados obtidos com o algoritmo LMS;
- Capítulo 6: Neste capítulo é apresentada a conclusão que resume os resultados obtidos no trabalho. Há ainda a presença de sugestões para propostas de trabalhos futuros.

# Capítulo 2

## Sistemas MIMO

*Este capítulo abordará os principais conceitos em relação aos sistemas MIMO de comunicação, que são fundamentais para o desenvolvimento deste trabalho*

### 2.1 Introdução

Este capítulo abordará os principais conceitos em relação aos sistemas MIMO de comunicação. Primeiramente, será mostrada a classificação de um sistema de acordo com o número de antenas no transmissor e no receptor, a partir disso, é possível a definição de quais sistemas podem ser classificados como MIMO. Além disso, será mostrado conceitos como diversidade espacial e multiplexação espacial, que são técnicas que exploram os caminhos multipercursos produzidos pelos sistemas MIMO.

Em sequência, será tratado o conceito de desvanecimento, que afeta os sistemas de comunicação sem fio e é responsável pela degradação do desempenho desse tipo de transmissão. É mostrada ainda a classificação dos sistemas tendo como base certos parâmetros como atraso por espalhamento, banda de coerência, desvio Doppler e tempo de coerência.

Por fim, é apresentado um modelo que descreve canais não correlacionados em sistemas MIMO. E, por fim, é apresentada uma forma de inserir a correlação entre os canais de comunicação no modelo de forma a aproximar esse ainda mais ao que é observado na prática.

### 2.2 Classificação dos Sistemas MIMO

O termo MIMO quando usado em seu sentido estrito é usado para sistemas com múltiplas antenas tanto no transmissor quanto no receptor. No entanto, esse termo pode ser usado também para sistemas que possuem múltiplas antenas no transmissor e apenas uma no receptor, *Multiple Input Single Output (MISO)* e para o caso em que há apenas uma antena no transmissor e múltiplas no receptor, *Single Input Multiple Output (SIMO)*. Dessa forma, esse termo não se aplica apenas a sistema com uma antena no transmissor e receptor - *Single Input Single Output (SISO)* Fig. 2.1 apresenta as configurações possíveis e as respectivas nomenclaturas, em que Tx representa o

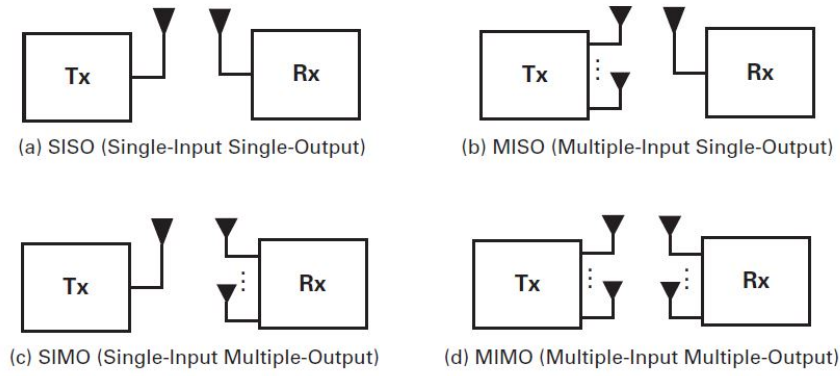


Figura 2.1: Possíveis configurações de antenas e nomenclatura atribuídas ao sistema. Adaptado de [2].

transmissor e Rx, o receptor. Ressalta-se que este capítulo tem como base a teoria exposta em [2].

Nesse contexto, é fundamental a análise do comportamento do sistema com o aumento do número de antenas. Essa ação simplesmente, é capaz de aumentar a capacidade do canal de forma subótima, é preciso, portanto, uma estimação do canal de comunicação para que a capacidade possa ser ampliada de forma otimizada. Para tanto, são necessários mecanismos adaptativos que diante desse cenário de variações temporais dos canais permitam a aproximação do comportamento real do canal de modelos matemáticos.

## 2.3 Diversidade Espacial e Multiplexação Espacial

Sistemas de comunicação sem fio estão sujeitos ao fenômeno de desvanecimento, que é responsável pela variação da intensidade do sinal recebido no tempo. Esse fenômeno é responsável pelo aumento da probabilidade de erro de *bit*, *bit error rate* (*BER*), para valores acima do que se verifica quando há apenas o ruído branco presente.

Uma das formas de se reduzir o impacto desse fenômeno é através da utilização de técnicas de diversidade, em que diversas cópias são transmitidas de forma a sofrerem o processo de desvanecimento de maneira independente. Assim, a medida em que se aumenta o número de cópias torna-se cada vez mais improvável que todas sejam atenuadas simultaneamente. E, através da combinação dessas no receptor é possível então se obter um sinal com melhor relação sinal ruído. Com isso, alcança-se também uma maior robustez do sistema e confiabilidade do sistema. É interessante, então, se analisar o número máximo de réplicas independentes de um sinal transmitido que podem ser usados no receptor para a combinação, isso é denominados ordem de diversidade ( $N_d$ ), cuja fórmula de cálculo é mostrada em Eq. (2.1).

$$Max(N_d) = N_t \times N_r \quad (2.1)$$

Por outro lado, deve-se analisar também qual a quantidade máxima de canais multi-percurso

resultantes do aumento no número de antenas. Esses canais permitem a transmissão em paralelo de informações, o que resulta em uma maior capacidade de transmissão de informações. Diversas denominações são dadas a esses canais entre elas pode-se citar fluxos de dados e autocanais. A existência de canais multipercurso depende de um ambiente rico em espalhamento.

O fenômeno dos múltiplos percursos é resultante dos fenômenos de reflexão, refração, difração, absorção e espalhamento. Sem esses fenômenos haveria entre os diversos autocanais correlação, que é responsável pela degradação do desempenho de sistema MIMO. Ademais, é possível ainda o cálculo do número de canais multipercurso por meio de Eq. (2.2).

$$N_{\text{autocanais}} = \min(N_t, N_r) \quad (2.2)$$

Os diversos autocanais sofrem processo de desvanecimento causado por inúmeros fenômenos como distância e objetos ou obstáculos na trajetória de propagação entre transmissor e receptor. Faz-se necessário então uma análise de como esses fatores influenciam o desempenho de sistemas MIMO, além disso, diversas definições advêm desse modelo de propagação de ondas eletromagnéticas e são essenciais para o bom entendimento dos trabalhos.

## 2.4 Fenômenos em canais multipercurso

A propagação de ondas eletromagnéticas na faixa de radiofrequência entre transmissor e receptor pode ocorrer de duas maneiras de forma direta ou indireta. O primeiro caso ocorre quando a trajetória não sofre os processos de reflexão, refração, difração, absorção e espalhamento. Nesse caso, há visada direta, que é também denominada *line-of-sight (LOS)*. A propagação ocorre, portanto, no espaço livre cujo fenômeno mais importante que ocorre com essa componente é o de atenuação pela atmosfera devido a grande distância de propagação, isso é responsável pela diminuição da potência recebida pelo receptor.

Por outro lado, a propagação indireta decorrente dos fenômenos descritos anteriormente é também denominada *Non-line-of-sight (NLOS)* que ocorre, portanto, sem visada direta. Esses fenômenos são apresentados em Fig. 2.2 e mostram como são importantes para a criação dos múltiplos percursos.

Além disso, os diversos mecanismos de propagação são responsáveis por variações na potência recebida de um mesmo sinal. Percebe-se que o sinal recebido pode ser decomposto em três componentes principais, que são representadas em Fig. 2.3. A linha pontilhada decrescente é a componente do sinal que mostra que há uma redução da potência recebida com o aumento da distância. Essa componente é descrita matematicamente pelo que denomina a lei de propagação da potência. Além disso, há em Fig. 2.3 a presença de duas outras curvas que representam os fenômenos de desvanecimento de grande e de pequena escala. A linha contínua com variações suaves representa o desvanecimento de grande escala enquanto a curva contínua com grandes variações representa o desvanecimento de pequena escala.

O desvanecimento de grande escala é decorrente de grandes objetos na trajetória de propa-

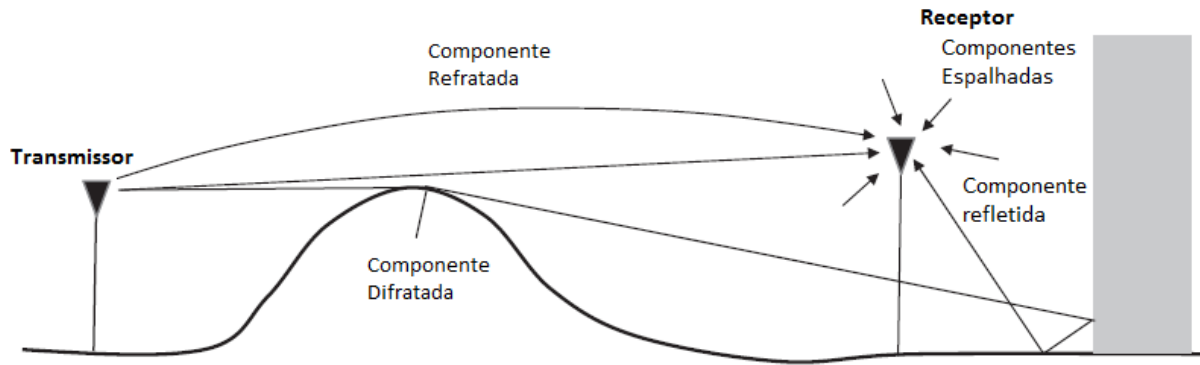


Figura 2.2: Fenômenos responsáveis por criar multi-percursos. Adaptado de [2].

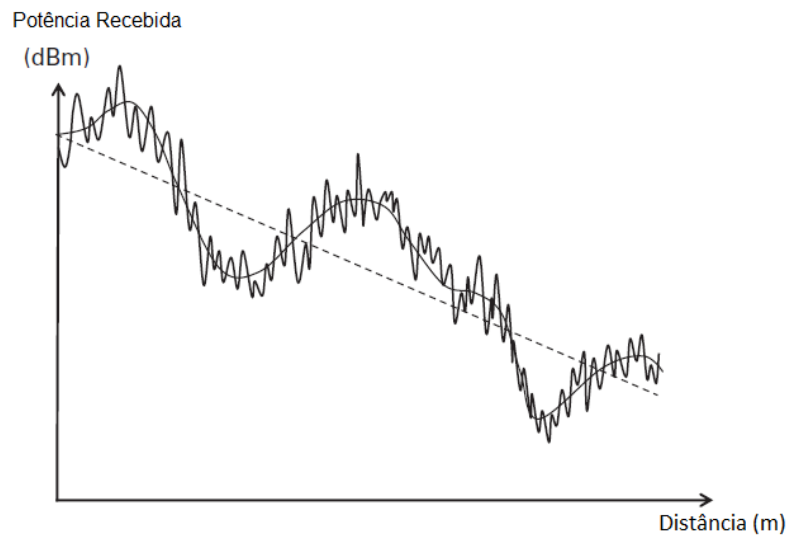


Figura 2.3: Representação do desvanecimento de grande escala e de pequena escala. Adaptado de [2].

gação, por isso, é chamado também de sombreamento. Por fim, o desvanecimento de pequena escala é uma características de canais multipercurso em que as variações da intensidade do sinal recebido são resultantes das interferências construtivas e destrutivas entre essas componentes resultantes dos variados mecanismos de propagação. A seguir são apresentados mais detalhes sobre o desvanecimento de grande e pequena escala.

## 2.5 Desvanecimento em grande escala

O desvanecimento em grande escala pode ser representado, então, como sendo composto de duas componentes uma decorrente da grande distância entre transmissor e receptor e outra componente resultante de grandes objetos no caminho de propagação do sinal.

A lei de propagação da potência possui diversas propriedades dependendo da natureza do canal de comunicação. Em geral, o cálculo da potência recebida pode ser feito usando a fórmula de

propagação de Friis, Eq. (2.3) - em que  $L_{prop}(d)$  são as perdas de propagação que são uma função da distância entre transmissor e receptor e  $L_{outras}$  são outras perdas causadas por diversas fontes como perdas em cabos, apontamento de antena e implementação. Além disso em Eq. (2.3),  $P_t$  é a potência transmitida,  $G_t$  é o ganho de transmissão e  $G_r$  é o ganho de recepção.

$$P_r = \frac{P_t G_t G_r}{L_{prop}(d) L_{outras}} \quad (2.3)$$

Uma das formas de se calcular as perdas de potência em função da distância é através da fórmula de perda no espaço livre, Eq. (2.4), que deve ser usada apenas quando os impactos de obstáculos na trajetória forem desprezíveis. Ademais, é necessário que haja visada direta entre transmissor e receptor. Em Eq. (2.4),  $d$  é a distância,  $f$  é a frequência do sistema e  $c$  é a velocidade da luz.

$$L_{fs}(d) = \left( \frac{4\pi df}{c} \right)^2 \quad (2.4)$$

Na fórmula da perda no espaço livre, percebe-se que a perda é proporcional ao inverso do quadrado da distância. Em Eq. 2.4, as perdas podem ter, no entanto, outro expoente dependendo do meio que propagam, podendo esse ser até mesmo igual a 20 em meios urbanos sem visada direta e, por outro lado, pode ser menor do que 2 em ambientes que guiam onda eletromagnética. Existem outras formas de cálculo da perdas em função da distância como o método dos dois raios que pode ser verificada em [2].

## 2.6 Desvanecimento em pequena escala

O desvanecimento de pequena escala é caracterizado por grandes mudanças na amplitude e fase decorrente de variações pequenas da ordem de comprimento de onda entre transmissor e receptor. Há dois mecanismos básicos de manifestação do desvanecimento de pequena escala: dispersão na frequência do sinal e variação temporal do canal.

A análise do desvanecimento de pequena escala é feita em dois domínios no tempo e na frequência. Por um lado, a dispersão do sinal pode ser dividida em dois tipos seletivo em frequência e não seletivo em frequência. Por outro lado, manifestações da variação temporal são classificadas em desvanecimento lento e desvanecimento rápido. Esses conceitos serão mais explorados nas seções seguintes com base nos conceitos desenvolvidos de parâmetros relacionados ao atraso e ao tempo.

Além disso, em um ambiente com múltiplos percursos em que o sinal recebido é composto por múltiplos raios e há a presença da componente de visada direta, componente especular, a envoltória apresenta uma função densidade de probabilidade Rician.

Quando não há a presença da componente especular, a envoltória apresenta um comportamento Rayleigh. A função de densidade de probabilidade Rayleigh é apresentada em Eq. (2.5) [13]. Nesse trabalho, o modelo de canal usado apresenta comportamento Rayleigh.



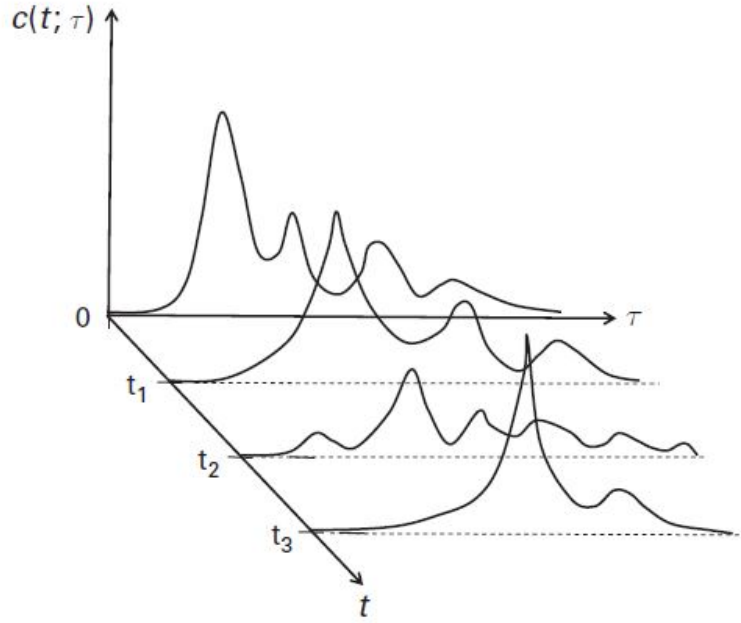


Figura 2.4: Representação da função resposta ao impulso de um sistema variante no tempo no domínio do tempo e do atraso. Adaptado de [2].

$$P(r) = \begin{cases} \frac{r}{\sigma} \exp \frac{r^2}{(2\sigma)^2}, & \text{para } r \geq 0 \\ 0, & \text{caso contrário} \end{cases} \quad (2.5)$$

Em Eq. (2.5),  $r$  é a amplitude da envoltória do sinal recebido e  $(2\sigma)^2$  é a média de potência predita do sinal.

Por fim, com base na introdução de um impulso no sistema, pode-se obter a resposta desse que pode ser descrita em um gráfico, que é apresentado em Fig. 2.4 cujos eixos representam o domínio do tempo e do atraso, apesar de na realidade tudo se dar no domínio do tempo. O eixo do tempo  $t$  representa o tempo em que o transmissor emite um impulso, além de ser a dimensão na qual o canal varia. Enquanto que o eixo  $\tau$  representa o tempo que o impulso leva para chegar ao receptor em que o tempo de emissão do impulso é a referência.

## 2.7 Parâmetros relacionados ao atraso

Existem dois parâmetros relacionados ao atraso para a classificação de um sistema de comunicação: espalhamento de atraso ( $T_m$ ) e banda de coerência ( $B_C$ ). Eles são, como visto anteriormente, relacionados a dispersão no tempo. Este fenômeno ocorre devido a presença dos múltiplos percursos. Esses conceitos, no entanto, não permitem um conhecimento da natureza do canal quanto a variação no domínio do tempo.

### 2.7.1 Espalhamento de atraso ( $T_m$ )

O espalhamento de atraso é uma medida da extensão do atraso de propagação enfreto, em um canal multipercurso variante no tempo, por suas componentes. A análise desse parâmetro tem como base a função de autocorrelação, que é usada para a definição do perfil de intensidade de múltiplos percursos,  $\phi_c(\tau)$ , também conhecido como espectro de atraso de potência.

O perfil de intensidade de múltiplos percursos pode ser medido na prática através da transmissão de um impulso e analisando a potência recebida como uma função do tempo. Em geral, quando há a componente especular o espectro de atraso de potência possui um pico logo nos instantes iniciais e em seguida decresce. Isso porque componentes que chegaram com maiores atrasos em relação a primeira componente percorreram maiores percursos e, portanto, foram mais desvanecidas em sua trajetória.

Pode-se ainda com base no atraso sofrido pelas diversas componentes se calcular o valor quadrático médio do máximo excesso de atraso,  $\sigma_\tau$ , que é uma das formas de se mensurar o espalhamento de atraso. Este pode ser calculado como mostrado em Eq. (2.6), em que  $\bar{\tau}$  é calculado como mostrado em Eq. (2.7) e  $\bar{\tau}^2$  como mostrado em Eq. (2.8).

$$\sigma_\tau = \sqrt{\bar{\tau}^2 - (\bar{\tau})^2} \quad (2.6)$$

$$\bar{\tau} = \frac{\sum_k \phi_c(\tau_k) \tau_k}{\sum_k \phi_c(t_k)} \quad (2.7)$$

$$\bar{\tau}^2 = \frac{\sum_k \phi_c(\tau_k) \tau_k^2}{\sum_k \phi_c(t_k)} \quad (2.8)$$

Em Fig. 2.5, é mostrado um perfil de atraso de potência, cujo excesso de atraso médio é calculado em Eq. (2.9), em que os valores absolutos foram substituídos por valores relativos. Por isso, as componentes que chegam com potência de -103 dBm são consideradas iguais a 0,5 pois possuem 50 % da potência da componente considerada como referência,  $\tau = 0$ .

$$\tau = \frac{(P)(0) + (0.5P)(100) + (0.5P)(200)}{P + 0.5P + 0.5P} \quad (2.9)$$

$$\tau = 75 \text{ ns} \quad (2.10)$$

Já o termo  $\bar{\tau}^2$  é calculado em Eq. (2.11).

$$\bar{\tau}^2 = \frac{(P)(0)^2 + (0.5P)(100^2) + (0.5P)(200^2)}{P + 0.5P + 0.5P} \quad (2.11)$$

$$\bar{\tau}^2 = 12500 \text{ ns}^2 \quad (2.12)$$

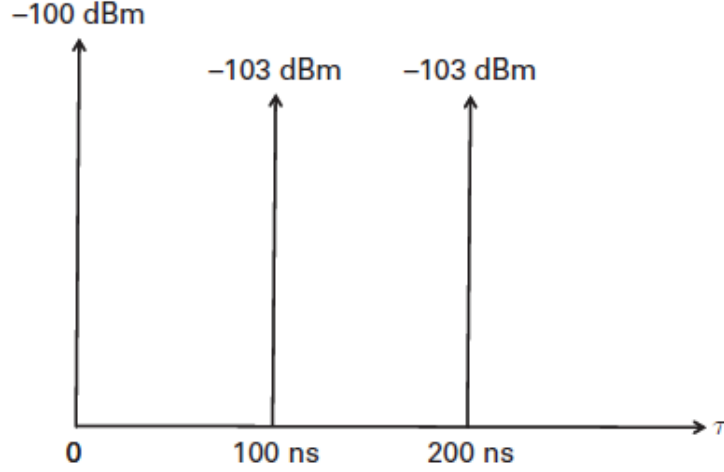


Figura 2.5: Espectro de atraso de potência. Adaptado de [2].

Com os resultados apresentados em Eq. (2.10) e em Eq. (2.12), pode-se então calcular o atraso rms,  $\sigma_\tau$ . Isso é feito em Eq. (2.13).

$$\sigma_\tau = \sqrt{12500 - (75)^2} \quad (2.13)$$

$$\sigma_\tau = 82,9 \text{ ns} \quad (2.14)$$

### 2.7.2 Banda de coerência ( $B_c$ )

A banda de coerência é uma extensão da faixa de frequências em que o canal é não distorcivo, assim, o canal permite a passagem de todas as componentes com igual ganho e atraso.

A banda de coerência e o atraso por espalhamento, como explicitado anteriormente são dois parâmetros relacionados ao atraso. A relação entre esses dois termos pode ser aproximada por uma relação matemática mostrada em Eq. (2.15). Essa equação é denominada banda de 50 %, em outras palavras, é a banda na qual a função de correlação no domínio da frequência -  $\Phi_c(\Delta f)$  - é maior do que 0,5.

$$B_{c,50} \approx \frac{1}{5\sigma_\tau} \quad (2.15)$$

A completa relação entre esses dois termos se dá através da transformada de Fourier. A relação matemática é mostrada em Eq. (2.16) e a ilustração dessa relação é mostrada em Fig. 2.6.

$$\Phi_c(\Delta f) = \int_{-\infty}^{\infty} \phi_c(\tau) e^{-2\pi\Delta f\tau} d\tau \quad (2.16)$$

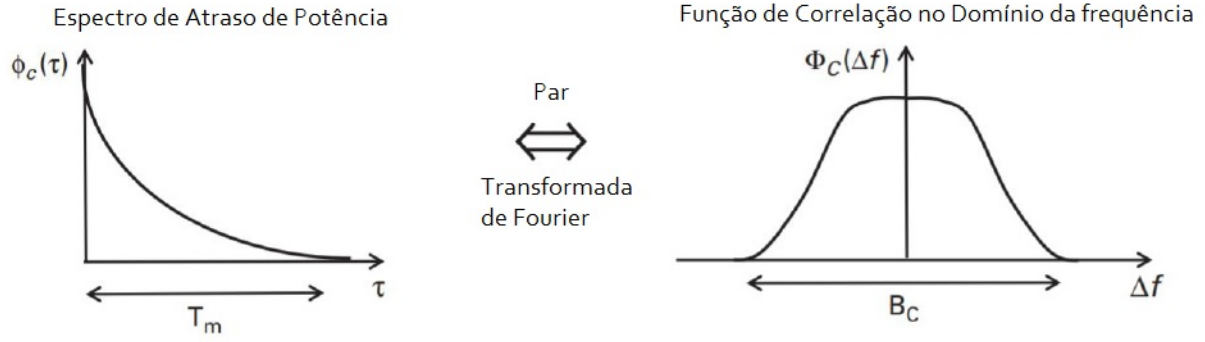


Figura 2.6: Par transformada de Fourier: atraso no tempo e banda de coerência. Adaptado de [2].

## 2.8 Parâmetros relacionados ao tempo

Os parâmetros relacionados ao tempo são o desvio Doppler,  $B_d$ , e o Tempo de coerência,  $T_c$ . Esses parâmetros são relacionados as características de dispersão na frequência do canal, que pode ser visto como sendo o espalhamento no domínio da frequência.

### 2.8.1 Desvio Doppler ( $B_d$ )

O espalhamento na frequência é um aumento da largura de banda ocupada por um sinal, como uma consequência do movimento relativo entre transmissor, espalhadores e receptor. Quando há o movimento relativo entre transmissor e receptor de forma a aproximá-los a uma velocidade,  $v$ , dada uma portadora de frequência,  $f_c$ , o receptor recebe uma frequência  $f_c + f_d$ , em que  $f_d$ , frequência Doppler, é mostrada em Eq. (2.18) e  $c$  é a velocidade da luz. O máximo desvio Doppler é definido então como mostrado em Eq. (2.17).

$$B_d = 2f_d \quad (2.17)$$

$$f_d = \frac{v f_c}{c} \quad (2.18)$$

### 2.8.2 Tempo de coerência ( $T_c$ )

O tempo de coerência está relacionado com a probabilidade de sinais serem altamente correlacionados dado um máximo atraso de transmissão máxima. Isso significa que se um canal varia lentamente o tempo de coerência será grande. O tempo de coerência e o espalhamento Doppler são inversamente proporcionais, dessa forma, quanto maior o espalhamento Doppler menor o tempo de coerência. Uma aproximação que pode ser feita é denominado valor de 50 %, isso significa que o tempo de coerência será medido até a correlação entre duas respostas impulsionais seja no mínimo 0,5.

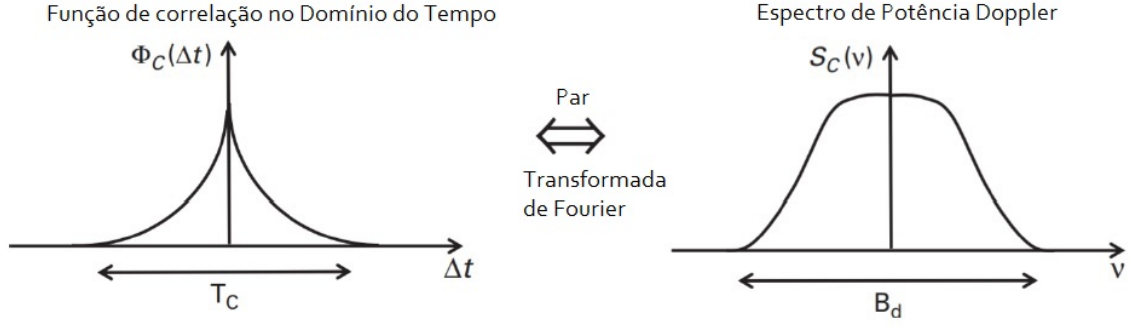


Figura 2.7: Par transformada de Fourier: espalhamento Doppler e tempo de coerência. Adaptado de [2].

$$T_{c,50} \approx \frac{9}{16\pi B_d} \quad (2.19)$$

De maneira análoga ao caso do domínio do atraso existe uma relação entre o espalhamento Doppler ( $B_d$ ) e o tempo de coerência ( $T_c$ ), que é descrita de maneira completa através da transformada de Fourier como mostrado em Eq. (2.20), em que  $S_c(\nu)$  é o espectro doppler e  $\Phi_c(\Delta t)$  é a função de correlação no domínio do tempo. Além disso, esse par transformada de Fourier é ilustrado em Fig. 2.7.

$$S_c(\nu) = \int_{-\infty}^{\infty} \Phi_c(\Delta t) e^{-2\pi\nu\Delta t} d\Delta t \quad (2.20)$$

### 2.8.3 Classificação dos canais multipercurso

Com base nos parâmetros no domínio do tempo e do atraso, os canais podem então ser classificados em desvanecimento plano ou seletivo, lento ou rápido.

O desvanecimento plano ou não-seletivo em frequência ocorre quando a largura de banda de um sinal,  $W$ , é menor do que a banda de coerência,  $B_c$ . Essa relação também pode ser vista como uma função do tempo de símbolo ( $T_c$ ) e do atraso por espalhamento ( $T_m$ ). Caso o tempo entre a transmissão de cada símbolo seja maior do que o atraso por espalhamento, o que significa que outro símbolo só será emitido após a chegada do símbolo anterior espalhado, dessa forma, sem causar interferência entre símbolos adjacentes, o canal será de desvanecimento plano. Para este tipo de canal a passagem de um símbolo com banda menor do que banda de coerência terá o efeito de apenas multiplicar todo o sinal por um ganho complexo.

Um canal seletivo em frequência é aquele em que a banda do sinal é maior do que a banda de coerência do canal. Isso significa que o espalhamento temporal de um sinal fará com que haja interferência intersimbólica. A transmissão do sinal por esse tipo de canal não é simplesmente a multiplicação deste por um ganho complexo, mas uma convolução da resposta ao impulso do canal com o sinal que foi introduzido na entrada do sistema.

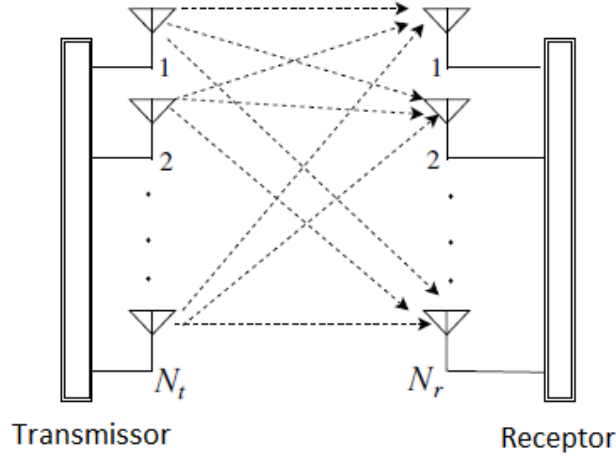


Figura 2.8: Modelo de sistema MIMO. Adaptado de [3].

A classificação de um canal como sendo de desvanecimento lento ou rápido está relacionado com a taxa de variação do canal. Essa é medida com base no tempo de coerência e no tempo de símbolo. Se o tempo de coerência for maior do que o tempo de símbolo, o canal permanecerá praticamente constante quando um símbolo for transmitido.

## 2.9 Modelo do Sistema

O efeito da transmissão de uma informação por um canal de comunicação com desvanecimento plano pode ser visto como sendo a multiplicação do sinal por um número completo, denominado resposta do canal. Além disso, em um sistema MIMO o sinal recebido em uma dada antena receptora será a soma dos sinais transmitidos por cada uma das antenas transmissoras como mostrado em Fig. 2.8. Assim, o sinal recebido pode ser escrito como em Eq. (2.21) em que  $z_i$  é o ruído da  $i$ -ésima antena.

$$r_i = \sum_{j=1}^{N_t} h_{ij} s_j + z_i, \quad i = 1, \dots, N_r \quad (2.21)$$

O sinal recebido pode ser escrito ainda de forma matricial como em Eq. (2.22). As indicações em negrito serão usadas para a indicação de matrizes.

$$\mathbf{r} = \mathbf{H}\mathbf{s} + \mathbf{z} \quad (2.22)$$

Em que,

$$\mathbf{r} = [r_1, \dots, r_{N_r}]^T \quad (2.23)$$

$$\mathbf{s} = [s_1, \dots, s_{N_t}]^T \quad (2.24)$$

$$\mathbf{z} = [z_1, \dots, z_{N_t}]^T \quad (2.25)$$

$$\mathbf{H} = \begin{pmatrix} h_{1,1} & \dots & h_{1,N_t} \\ \vdots & \ddots & \vdots \\ h_{N_r,1} & \dots & h_{N_r,N_t} \end{pmatrix} \quad (2.26)$$

Em Eq. (2.22),  $r_i$  é o sinal recebido na  $i$ -ésima antena receptora,  $s_j$  é o sinal transmitido pela  $j$ -ésima antena transmissora,  $h_{i,j}$  são os ganhos, que é a resposta do canal entre a  $i$ -ésima antena transmissora e  $j$ -ésima antena receptora e  $z$  é o ruído da  $i$ -ésima antena receptora.

Como visto em [2], a capacidade de um sistema MIMO em que são usadas as componentes em fase e em quadratura pode ser aproximada por uso de canal como em Eq. (2.27). Percebe-se ainda a utilização do logaritmo na base 2, assim, a capacidade terá a unidade de *bits*.

$$C_{complexo} = \log_2 \left| \mathbf{I}_{N_r} + \frac{1}{\sigma^2} \mathbf{H} \mathbf{R}_{ss} \mathbf{H}^H \right| \quad (2.27)$$

Para a condição em que apenas o receptor tem informações sobre a matriz de ganhos do canal, CSIR, a potência é dividida no transmissor entre todos os autocanais de forma idêntica. Normalizando a energia que é colocada nos autocanais, tem-se a relação Eq. (2.28).

$$E[|s_i|^2] = \sigma_s^2 = 1/N_t \quad (2.28)$$

Muitas vezes, é interessante a análise da capacidade de canais MIMO tendo como base a relação sinal-ruído,  $\rho$ . Para que ela seja explicitada é preciso, então, considerar certas condições de normalização mostradas em Eq. (2.29), Eq. (2.30) e Eq. (2.31). A primeira condição de normalização representa que o ganho absoluto médio do sistema MIMO é igual a 1, já a segunda equação de normalização significa que a energia média do ruído é igual a 1. Por fim, a última equação de normalização significa que a potência de transmissão é igualmente dividida entre as antenas transmissoras. Com base nessas equações, pode-se, então escrever o sinal recebido como mostrado em Eq. (2.32). Ressalta-se ainda que  $h'_{i,j}$  representa os coeficientes da matriz de ganhos normalizada enquanto que  $h_{i,j}$  representa os termos da matriz de ganho não normalizada. Para simplificação de notação, a partir desse ponto, o apóstrofo será suprimido apesar de se tratar dos coeficientes referentes a matriz de ganhos normalizada.

$$E[|h'_{i,j}|^2] = 1, \quad (2.29)$$

$$E[|z_i|^2] = \sigma^2 = 1, \quad (2.30)$$

$$E[|s_i|^2] = \sigma_s^2 = \frac{1}{N_t} \quad (2.31)$$

$$\mathbf{r} = \sqrt{\rho}\mathbf{H}'\mathbf{s} + \mathbf{z} \quad (2.32)$$

Com base em Eq. (2.27), e considerando Eq. (2.28) pode-se escrever então a capacidade como mostrada em Eq. (2.33).

$$C = \log_2 |\mathbf{I}_{Nr} + \frac{\sigma_s}{\sigma^2} \mathbf{H}\mathbf{H}^H| \quad (2.33)$$

E, a partir das normalização presentes em Eq. (2.30) e Eq. (2.31) pode-se chegar em Eq. (2.34).

$$C = \log_2 |\mathbf{I}_{Nr} + \frac{\rho}{N_t} \mathbf{H}\mathbf{H}^H| \quad (2.34)$$

Uma vez que  $\mathbf{H}\mathbf{H}^H$  em Eq. (2.34) é uma matriz Hermitiana, pode-se fazer a decomposição desse produto de forma a se evidenciar os autovalores desse produto. Isso é feito através do produto de matrizes  $\mathbf{U}\mathbf{D}\mathbf{U}^H$  em que esses autovalores estão presentes na diagonal principal de  $\mathbf{D}$  e  $\mathbf{U}$  é uma matriz identidade. Através da manipulação de Eq. (2.35) pode-se chegar em uma fórmula de capacidade de sistema MIMO em que são explicitados os valores do autocanais, Eq. (2.40). Nesta equação,  $r$  representa o posto do canal de comunicação que é igual ao  $\min(N_t, N_r)$ .

$$C = \log_2 |\mathbf{I}_{Nr} + \frac{\rho}{N_t} \mathbf{U}\mathbf{D}\mathbf{U}^H| \quad (2.35)$$

$$C = \log_2 |\mathbf{U}(\mathbf{I}_{Nr} + \frac{\rho}{N_t} \mathbf{D})\mathbf{U}^H| \quad (2.36)$$

$$C = \log_2 [\det(\mathbf{U}).\det(\mathbf{I}_{Nr} + \frac{\rho}{N_t} \mathbf{D}).\det(\mathbf{U}^H)] \quad (2.37)$$

$$C = \log_2 |(\mathbf{I}_{Nr} + \frac{\rho}{N_t} \mathbf{D})| \quad (2.38)$$

$$C = \log_2 [\prod_{i=1}^r (1 + \frac{\rho}{N_t} \lambda_i)] \quad (2.39)$$

$$C = \sum_{i=1}^r \log_2 (1 + \frac{\rho}{N_t} \lambda_i) \quad (2.40)$$

É natural então se analisar, de acordo com os autovalores dos canais, qual distribuição permite a capacidade ótima quando apenas o receptor tem informações sobre a matriz de ganho. Isso foi feito como mostrado em Eq. (2.9). Nessa simulação gerada no MATLAB, foram feitas 1000 iterações em que foram gerados 4 autovalores aleatórios, cuja soma total era igual a 1. Percebe-se, a partir dos resultados, que a máxima capacidade é atingida para o casos de potência igualmente dividida



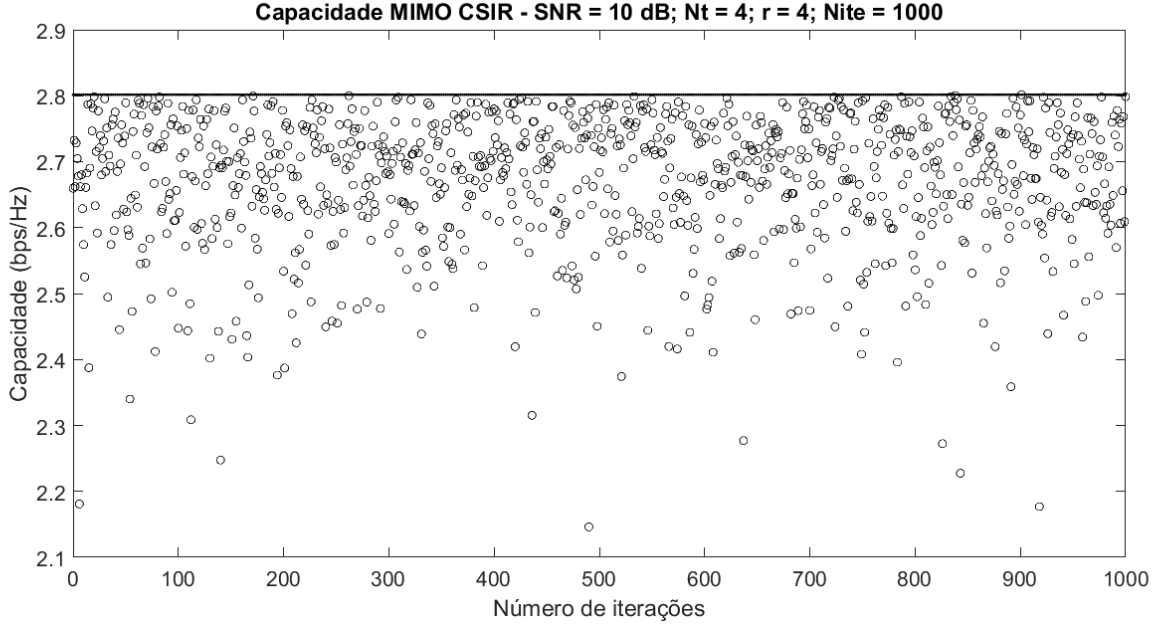


Figura 2.9: Capacidade ótima para sistema MIMO  $4 \times 4$ , SNR = 10 dB e 1000 realizações.

entre os autovalores. Tudo isso na situação em que apenas o receptor tem informações sobre a matriz de ganhos do sistema. Essa capacidade é representada pela linha em aproximadamente 2,8 bps/Hz. Ressalta-se que essa não é a máxima capacidade possível de um sistema, mas a máxima possível quando apenas o receptor tem informações sobre a matriz de ganhos.

Anteriormente, no caso CSIR, a potência era igualmente distribuída entre os autovalores. No caso em que ambos envolvidos no processo de comunicação, transmissor e receptor, têm informações sobre a matriz de ganhos do sistema, pode-se então otimizar a distribuição de potência. Através da conformação de feixes (*eigenbeamforming*) é possível se escolher de forma apropriada a potência para cada um dos autocanais.

Uma das formas de se escolher de forma apropriada a potência que deve ser empregada para cada um dos autocanais é através da utilização do algoritmo de waterfilling. Esse algoritmo tem como base um problema de otimização envolvendo a técnica dos multiplicadores de Lagrange, cuja função de custos a ser maximizada é mostrada em Eq. (2.41).

$$J(P_1, P_2, \dots, P_r) = \sum_{i=1}^r \log_2(1 + \rho P_i \lambda_i) + \beta \sum_{i=1}^r P_i \quad (2.41)$$

Com base em um sistema MIMO com 3 antenas transmissoras, 4 antenas receptoras e, portanto, posto 3, e cuja matriz de ganho é mostrada em Eq. (2.42), foi feita a utilização do algoritmo de waterfilling. Com base na execução dessa simulação, tem-se como objetivo a verificação do processo de otimização da distribuição de potência entre os diversos 3 autocanais presentes nesse sistema.

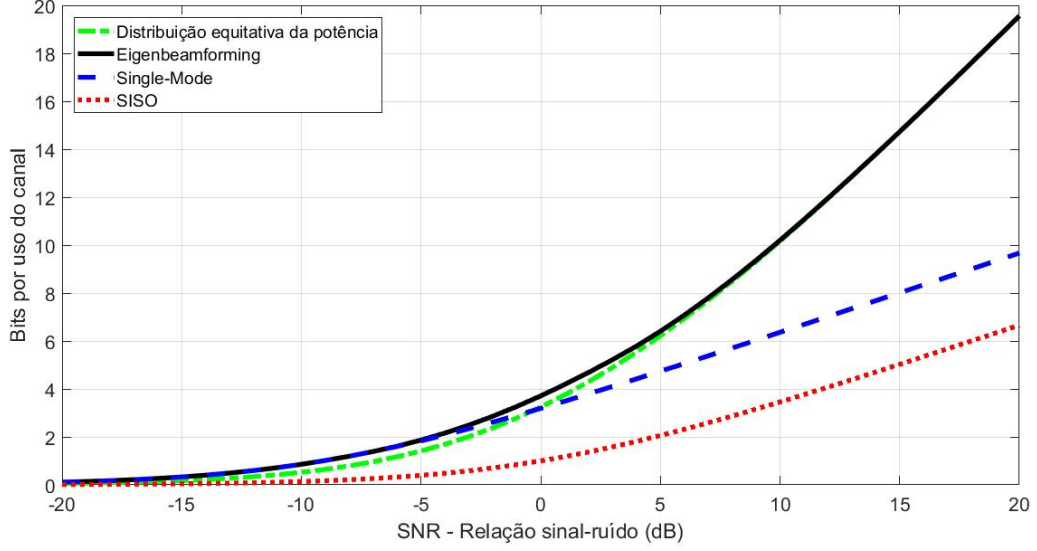


Figura 2.10: Comparações entre as capacidades do canal para matriz de ganhos descrita em Eq. (2.42).

$$\mathbf{H}_1 = \begin{pmatrix} 0.460 + j0.31 & -0.079 + j0.18 & 0.43 - j0.68 \\ -0.040 + j0.66 & 0.950 - j0.50 & -0.77 - j0.51 \\ 0.450 + j0.46 & 0.900 + j0.56 & 0.46 - j1.87 \\ -0.013 - j0.19 & 0.900 - j0.73 & 1.04 + j0.91 \end{pmatrix} \quad (2.42)$$

Com base em Fig. 2.10, percebe-se que para valores de SNR baixos a conformação de feixes (*eigenbeamforming*) permite maior capacidade para o canal comparado a qualquer um dos outros esquemas. E, como descrito anteriormente, para que esse processo ocorra é necessário que o transmissor tenha informações da matriz de ganhos. Além disso, é necessária a existência de um sistema em malha fechada que permita o transmissor ter informações sobre a matriz de ganhos. Por fim, percebe-se ainda que a aumentando a SNR, potência é igualmente distribuída entre os diversos autocanais.

Ademais, a técnica em que se utiliza o autocanal cujo autovalor é o maior possível é denominada *single-mode*, que possui ainda ganhos em relação ao caso SISO. Para que esse método possa ser utilizado é necessária também a estimação dos canais de comunicação de maneira adequada. Nota-se ainda que todos os esquemas MIMO possuem desempenho superiores ao caso SISO.

Portanto, mostra-se, com isso, que é essencial a correta estimação dos canais de comunicação para que os ganhos com sistemas MIMO sejam ampliados. Por isso, faz-se necessário algoritmos que permitam uma melhor estimação dos canais de comunicação como o filtro de Kalman, cujos conceito e princípios serão melhor descritos no capítulo 5, e que apresenta desempenho superior a outros métodos de estimação de canal como o algoritmo LMS, apresentado em no capítulo 4.

## 2.10 Modelo de Sistema com Correlação

Até o momento foi considerado que os canais são não-correlacionados. No entanto, na realidade existe um certo grau de correlação que é responsável pela degradação de parte da capacidade do canal de comunicação.

Essa correlação pode ser calculada com base em Eq. (2.43), em que  $vec$  é um vetor coluna formado pela concatenação das colunas de  $\mathbf{H}$ . Com esse conceito é possível o cálculo entre a correlação entre todos os possíveis canais.

$$\mathbf{R} = E[vec(\mathbf{H})vec(\mathbf{H}^H)] \quad (2.43)$$

No entanto, o método empregado para cálculo como visto em Eq. (2.43) pode exigir grande esforço computacional, que pode ser simplificado através da utilização do modelo de Kronecker de canal. Nesse modelo a correlação total do sistema entre os autocanais pode ser subdivida entre a correlação inserida pelas antenas transmissoras e a correlação inserida pelas antenas receptoras. Dessa forma, pode-se por meio do produto entre a matriz de correlação nas antenas transmissoras, a matriz de ganhos de um sistema não correlacionado e da matriz de correlação das antenas receptoras obter a matriz de ganhos de um sistema correlacionado.

$$\mathbf{H} = \mathbf{R}_r^{(1/2)} \mathbf{G} \mathbf{R}_t^{(1/2)} \quad (2.44)$$

Em Eq. (2.44),  $\mathbf{R}_r$ ,  $\mathbf{G}$  e  $\mathbf{R}_t$  são, respectivamente, a matriz de correlação entre as antenas receptora, a matriz de ganhos do sistema sem correlação e a matriz de correlação entre as antenas transmissoras.

## 2.11 Conclusões parciais

Nestes capítulo foi feita a classificação de sistema de comunicação com base no número de antenas que estão presentes no receptor e no transmissor. Essa classificação foi feita com base em uma denominação em sentido amplo de sistema MIMO.

Em seguida, foram abordados conceitos importantes dentro de sistema MIMO, que são os de diversidade espacial e multiplexação. Esses são altamente explorados dentro desse tipo de sistema e, permitem com que sejam alcançados melhores desempenhos, seja em capacidade, seja em robustez do sistema.

Foram tratados ainda os conceitos de desvanecimento em grande escala e pequena escala que são descrições de fatores que afetam os sistema de comunicação sem fio. E, foi descrito como cada um desses fatores são responsáveis pela degradação do desempenho do sistema.

A descrição de parâmetros que permitem a classificação de sistemas MIMO também foi feita. Estes parâmetro foram classificados em relacionados ao atraso e ao tempo. No primeiro grupo,

estão o espalhamento de atraso ( $T_m$ ) e a banda de coerência ( $B_c$ ), enquanto no segundo grupo estão o desvio Doppler ( $B_d$ ) e o tempo de coerência ( $T_c$ ).

Por fim, foi apresentado um modelo de sistema MIMO sem correlação entre os autocanais, que pode ser descrito como um sistema matricial. Foi analisada a capacidade desse sistema e comparado o desempenho de sistemas MIMO em relação aos sistemas SISO. E, como uma maneira de aproximar o modelo MIMO anteriormente descrito ao que geralmente ocorre foi analisado como se inserir a correlação nesse modelo de sistema MIMO.

## Capítulo 3

# Modelo de Canal

*Este capítulo fará uma descrição do modelo de canal, que será usado para ser rastreado por meio do algoritmo LMS e do filtro de Kalman que serão descritos nos capítulos 4 e 5.*

### 3.1 Introdução

Neste capítulo, há uma breve descrição do modelo de canal de Clark em que são apresentadas suas semelhanças e diferenças em relação ao canal de Jakes, que foi usado neste trabalho para o rastreamento do canal. O canal de Jakes foi de extrema importância, uma vez que foi o primeiro modelo de canal de comunicações móveis, a permitir também a análise do efeito Doppler e os desvanecimentos de amplitude [4].

Por conseguinte, há a descrição matemática do canal de Jakes, essencial para se compreender de maneira completa o modelo e se avaliar posteriormente a correta implementação do canal. Há também uma descrição da implementação desse canal partindo de um resultado da teoria de processos estocásticos. Por fim, há a apresentação de diversas simulações que visam corroborar a consistência do modelo implementado e seu casamento com a teoria que o descreve.

### 3.2 Desvio Doppler e modelo de canal de Clark

O desvio Doppler, como descrito no capítulo 2, é causado pelo movimento relativo entre transmissor, espalhadores e receptor. Esse fenômeno é descrito em Eq. (2.17). No caso em que se tem uma portadora de frequência,  $f_c$ , e caso o movimento relativo entre transmissor e receptor seja de aproximação, cria-se a componente de frequência  $f_c + f_d$  em que  $f_d$  é descrito como em Eq. (3.1). Por outro lado, quando há um movimento de separação entre esses, há a criação de componente cuja frequência é de  $f_c - f_d$ . Isso é ilustrado em Fig. 3.1. Assim como mostrado nesta figura, há um alargamento do espectro recebido, que é denominado desvio Doppler.

$$f(d) = \frac{v f_c}{c} \quad (3.1)$$

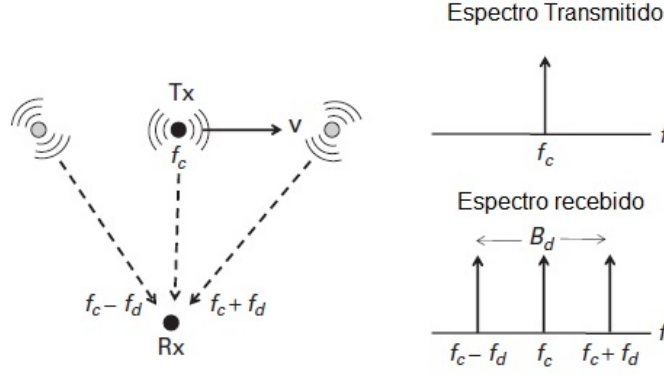


Figura 3.1: Representação do desvio Doppler. Adaptado de [2].

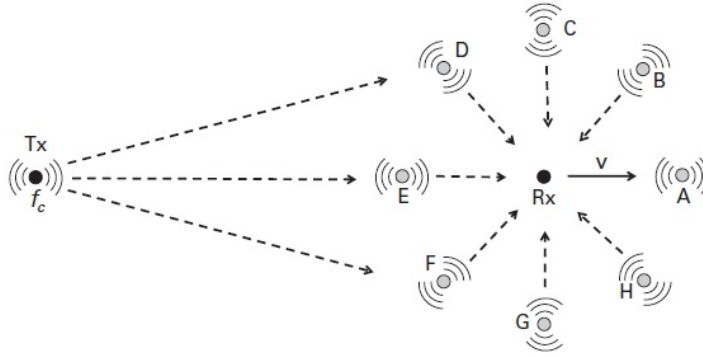


Figura 3.2: Representação do modelo de canal de Clark [2].

Diversos modelos são desenvolvidos com o objetivo de retratar ambientes reais para que se possa realizar a simulação do comportamento desses ambientes. Um desses modelos é o de Clark, que permite também a análise do desvio Doppler e tenta retratar ambientes urbanos em que o sinal que chega ao receptor advém do espalhamento em construções e objetos em volta do receptor.

Nesse modelo, o transmissor encontra-se fixo e o sinal recebido vem do espalhamento de infinitos espalhadores em um arranjo circular em volta do receptor, que se move. As componentes de frequência recebidas nesse modelo estarão entre os extremos descritos anteriormente de  $f_c - f_d$  e  $f_c + f_d$ . Além disso, não existe a componente especular. A representação desse modelo de Clark é mostrada em Fig. 3.2.

O espectro recebido,  $S(f)$ , no modelo de Clark foi analisado por Gans [16] para antenas receptoras verticais de comprimento  $\lambda/4$  e pode ser aproximado como mostrado em Eq. (3.2). Ademais, em Fig. 3.3, foi feita uma simulação no MATLAB em que são apresentados os espectro de potência recebido para diferentes velocidades do receptor. Percebe-se que o formato em  $U$  é típico desse modelo de canal, além disso quando o espalhador é perpendicular a trajetória do receptor o espalhamento é zero.

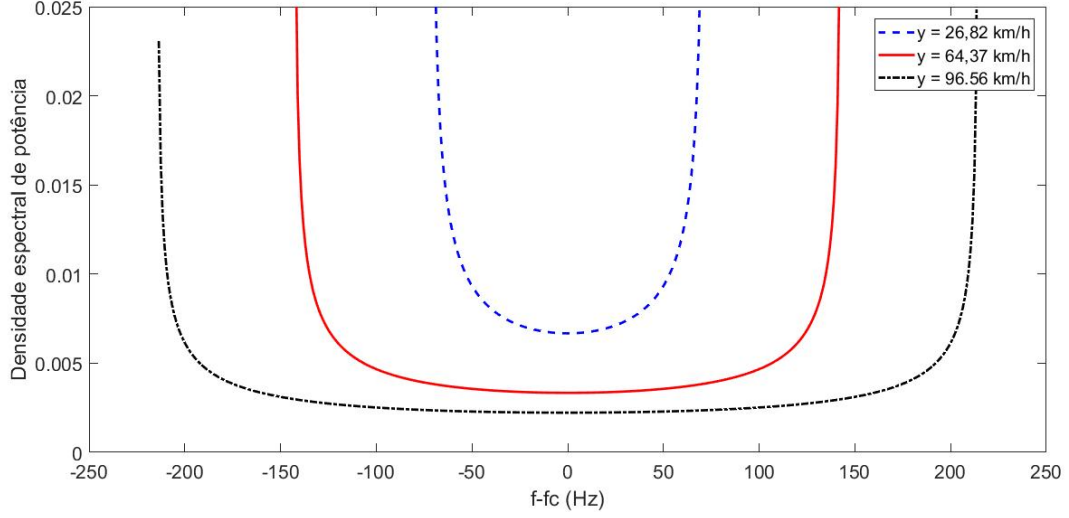


Figura 3.3: Espectro de Potência Doppler para portadora de frequência igual a 2,4 GHz.

$$S(f) = \frac{1.5}{\pi f_d \sqrt{1 - \left(\frac{f-f_c}{f_d}\right)^2}} \quad (3.2)$$

Uma vez que não há visada direta entre o transmissor e o receptor, como descrito no capítulo 2, o modelo de Clark é então um modelo cuja função densidade de probabilidade é Rayleigh. A geração desse modelo de canal pode ser feita através da utilização de duas distribuições Gaussianas. Assim, caso a resposta do canal seja descrita por um número complexo, a componente em fase,  $r_c(t)$ , e em quadratura,  $r_s(t)$ , devem ser gaussianas e independentes para que a envoltória do canal,  $r(t)$ , mostrada em Eq. (3.3) seja descrita por uma função de densidade de probabilidade Rayleigh.

$$r(t) = \sqrt{r_c(t)^2 + r_s(t)^2} \quad (3.3)$$

### 3.3 Modelo de canal de Jakes

O modelo de canal de Jakes foi usado por Jakes e outros nos laboratórios Bell para se chegar ao primeiro modelo de canal de comunicações móveis para análise do efeito Doppler e desvanecimentos na amplitude.

O modelo de canal de Jakes é similar ao modelo de Clark em que o sinal recebido é descrito como sendo uma somatória de componentes, mas nesse caso senoidais. Além disso, também considera-se que o espalhamento seja isotrópico, vindos de todas as direções. Nesse modelo, no entanto, o ângulo de chegada de cada componente é separado de  $2\pi/N$  radianos, em que  $N$  é o número de componentes. Cada componente que chega ao receptor possui diferentes atrasos, porém todas tem o mesmo espectro Doppler [5].

### 3.4 Desenvolvimento matemático

Pode-se representar o sinal no receptor como uma superposição de ondas planas, cuja envoltória complexa é mostrada em Eq. (3.4) [4]. Nesta equação,  $w_c$  é igual a  $2\pi f_c$  em que  $f_c$  é igual a frequência da portadora.

$$E(t) = \text{Re}[T(t)e^{iw_ct}] \quad (3.4)$$

em que,

$$T(t) = E_o \sum_{n=1}^N c_n e^{(iw_m t \cos \alpha_n + \phi_n)} \quad (3.5)$$

Em Eq. (3.5), a função de distribuição de probabilidade,  $c_n$ , pode ser descrita ainda como em Eq. (3.8) em que a função de densidade de probabilidade,  $p(\alpha_n)$ , é igual a  $\frac{1}{2\pi}$ . Assume-se ainda que os ângulos de chegada são distribuídos de maneira uniforme, ou seja,  $d\alpha = 2\pi/N$ . Portanto, tem-se Eq. (3.6) e Eq. (3.7), em que  $N$  é o número de componentes que chegam ao receptor.

$$c_n^2 = \frac{1}{N} \quad (3.6)$$

$$\alpha_n = \frac{2\pi n}{N}, \quad \text{para } n = 1, 2, \dots, N \quad (3.7)$$

$$c_n^2 = p(\alpha_n) d\alpha = \frac{1}{2\pi} d\alpha \quad (3.8)$$

Além disso,  $w_m$  é a frequência Doppler de uma dada componente definida em função do ângulo  $\alpha_n$  entre a reta que liga o espalhador e o receptor, ilustrado em Fig. 3.4 e definido em Eq. (3.9) em função da máxima frequência Doppler em radianos,  $w_M$ , que por sua vez é definido em Eq. (3.10).

$$w_m = w_M \cos(\alpha_n) \quad (3.9)$$

$$w_M = 2\pi f_d \quad (3.10)$$

Rearranjando Eq. (3.5) e considerando que se tem simetria par no arranjo, chega-se a Eq. (3.11) em que o primeiro termo representa o desvio Doppler de  $+w_m \cos(2\pi/N)$  a  $-w_m \cos(2\pi/N)$ , o segundo termo representa o desvio Doppler porém dessa vez de  $-w_m \cos(2\pi/N)$  a  $+w_m \cos(2\pi/N)$ . Por fim, os dois últimos termos representam os máximos desvio Doppler positivos e negativos respectivamente. Percebe-se que em Eq. (3.11) há sobreposição dos desvios Doppler, esta pode, no entanto, ser reescrita sem que haja sobreposição como mostrado em Eq. (3.12), em que  $N_o = \left(\frac{N}{2} - 1\right)$  e o termo  $\sqrt{2}$  foi inserido para que não houvesse alteração na potência total.



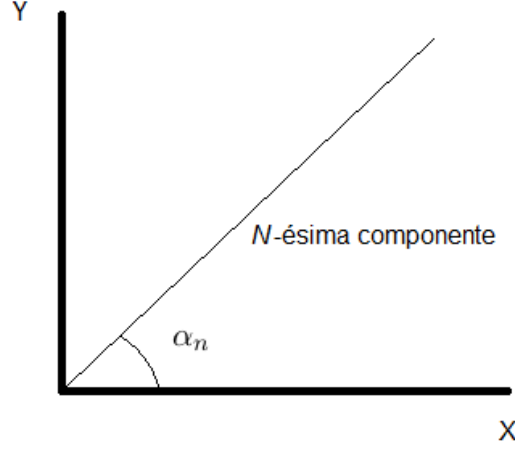


Figura 3.4: Ângulo de chegada da componente de um espalhador.

$$T(t) = \frac{E_o}{\sqrt{N}} \sum_{n=1}^{N/2-1} [e^{i(w_m t \cos \alpha_n + \phi_n)} + e^{-i(w_m t \cos \alpha_n + \phi_{-n})}] + e^{i(w_m t + \phi_N)} + e^{-i(w_m t + \phi_{-N})} \quad (3.11)$$

$$T(t) = \frac{E_o}{\sqrt{N}} \sqrt{2} \sum_{n=1}^{N_o} [e^{i(w_m t \cos \alpha_n + \phi_n)} + e^{-i(w_m t \cos \alpha_n + \phi_{-n})}] + e^{i(w_m t + \phi_N)} + e^{-i(w_m t + \phi_{-N})} \quad (3.12)$$

Com base em Eq. (3.12), em grandes valores de  $N$  e no Teorema central do limite pode-se concluir que  $T(t)$  é aproximadamente um processo gaussiano complexo, em que  $|T|$  é Rayleigh como desejado.

Analisa-se, então, a função de autocorrelação de  $T(t)$ , que é descrita matematicamente em Eq. (3.13), em que os colchetes angulares representam a média temporal. Esta é feita com base na esperança, que é tomada em relação aos termos  $\phi_n$  e  $\phi_N$ . Por fim, pode-se escrevê-la como em Eq. (3.14), em que os termos que contribuem para essa correlação são os casos em que  $n = N$ .

$$R(\tau) = \langle E(t) E^*(t + \tau) \rangle \quad (3.13)$$

$$R(\tau) = \frac{b_o}{N} \cos w_c \tau \left[ 4 \sum_{n=1}^{N_o} \cos \left( w_m \tau \cos \frac{2\pi n}{N} \right) + 2 \cos(w_m \tau) \right] \quad (3.14)$$

Analisando-se Eq. (3.14) percebe-se que a forma desta se assemelha a de uma portadora multiplicada por um tempo de baixa frequência como mostrado em Eq. (3.15)

$$R(\tau) = g(\tau) \cos(w_c \tau) \quad (3.15)$$

Jakes [4] mostra que se o espalhamento é uniforme, então,  $g(\tau)$  é descrito como em Eq. (3.16), em que  $J_o(x)$  é a função de Bessel de primeira espécie e ordem zero, definida como em Eq. (3.17).

$$g(\tau) = b_o J_o(w_m \tau) \quad (3.16)$$

$$J_o(x) = \frac{2}{\pi} \int_0^{\pi/2} \cos(x \cos \alpha) d\alpha \quad (3.17)$$

Analizando ainda o termo entre colchetes presente em Eq. (3.14) e fazendo a devida relação com Eq. (3.15), pode-se chegar na relação mostrada em Eq. (3.18). Nesta, há uma relação entre uma integral contínua em um somatório discreto quando  $N_o$  tende a um valor grande. Jakes [4] analisa ainda esta última relação em que a densidade espectral de potência é uma aproximação da relação presente em Eq. (3.19), que se aproxima, portanto, da relação presente em Eq. (3.2).

$$2 \left[ \sum_{n=1}^{N_o} \cos \left( w_m \tau \cos \frac{2\pi n}{N} \right) + \cos(w_m \tau) \right] = \frac{N}{2} J_o(w_m \tau) \quad (3.18)$$

$$S_n(f) = \left[ 1 - \left( \frac{f - f_c}{f_d} \right) \right]^{-1/2} \quad (3.19)$$

### 3.5 Construção do modelo

A construção do modelo pode ser feita então com base em  $N_o$  osciladores de baixa frequência, cujas frequências são definidas a partir dos desvios Doppler em Eq. (3.9). Outrossim, é necessário também outro oscilador cuja frequência é igual a  $w_M$ . As amplitudes de todas as componentes são feitas iguais a 1, exceto a da componente cuja frequência é igual a  $w_M$  em que a amplitude é definida como igual a  $1/\sqrt{2}$ .

As fases dessas componentes,  $\beta_n$ , são definidas de maneira que a distribuição de probabilidade seja a mais próxima da distribuição uniforme de  $1/2\pi$ . Jakes [4] mostra que alicerçado em relações trigonométricas, as fases podem ser geradas por amplificadores cujos ganhos são iguais a  $2\cos(\beta_n)$  ou  $2\sin(\beta_n)$ .

Cada componente é formada, então, pela saída dos osciladores com os ganhos respectivos para cada caso, que são somados para, então, formarem as componentes em fase,  $x_c$ , mostrado em Eq. (3.20) e em quadratura,  $x_s$ , mostrada em Eq. (3.21). Por conseguinte, são multiplicadas pelas componentes da portadora em fase e em quadratura e, por fim, somadas para produzir o sinal final,  $y(t)$ . A representação da construção do modelo é mostrada em Fig. 3.5.

$$x_c(t) = 2 \sum_{n=1}^{N_o} \cos \beta_n \cos w_n t + \sqrt{2} \cos \alpha \cos w_m t \quad (3.20)$$

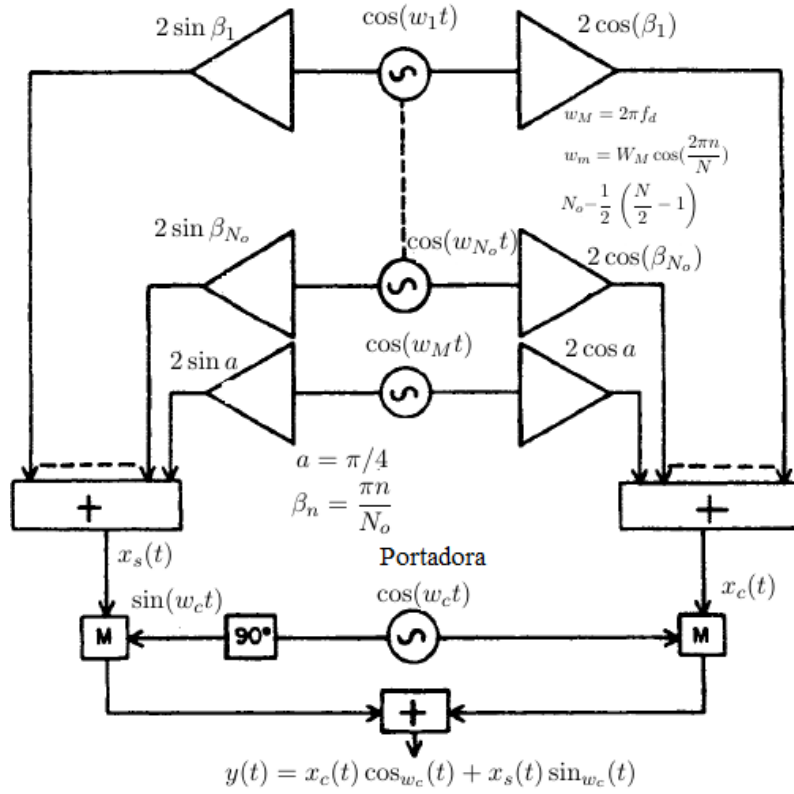


Figura 3.5: Representação da construção do modelo de Jakes. Adaptado de [4].

$$x_c(t) = 2 \sum_{n=1}^{N_o} \sin \beta_n \cos w_n t + \sqrt{2} \sin \alpha \cos w_m t \quad (3.21)$$

A fase  $y(t)$  deve ser aleatória e uniformemente distribuída no intervalo de 0 a  $2\pi$ . Isso faz com que  $\langle x_c^2 \rangle \approx \langle x_s^2 \rangle$  e  $\langle x_c x_s \rangle$ , em que os colchetes angulares representam média temporal.

Escolhendo-se  $a = 0$ ,  $\beta_n = \pi n / (N_o + 1)$ , chega-se em Eq. (3.22), Eq. (3.23) e Eq. (3.24). Com essas escolhas as componentes em fase e em quadratura são processos aleatórios gaussianos e, portanto, a envoltória,  $|y|$ , apresenta uma distribuição Rayleigh.

$$\langle x_c x_s \rangle \equiv 0 \quad (3.22)$$

$$\langle x_c^2 \rangle = N_o \quad (3.23)$$

$$\langle x_s^2 \rangle = N_o + 1 \quad (3.24)$$

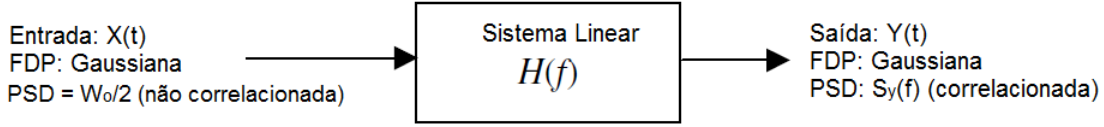


Figura 3.6: Descrição do processo necessário para a criação de uma sequência aleatória com uma dada correlação. Adaptado de [5].

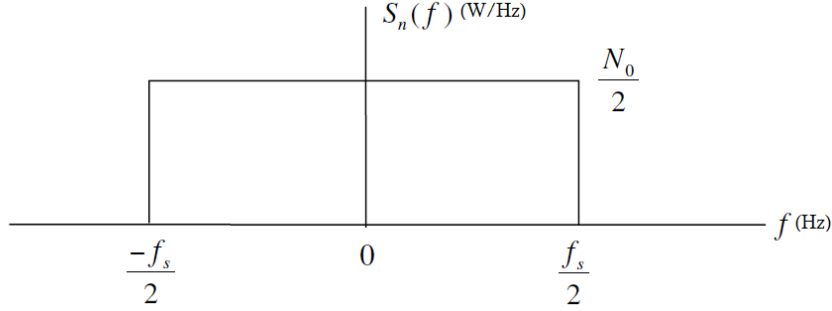


Figura 3.7: Densidade espectral de potência para amostras independentes. Adaptado de [5].

### 3.6 Modelagem do canal por meio da conformação espectral

A construção do modelo de canal de Jakes de forma tradicional não garante que as realizações sejam não correlacionadas como desejado. Deve-se então buscar novas abordagens que permitam a construção desse modelo de canal, uma dessas é mostrada neste trabalho por meio da utilização da ideia de filtro [5]. Para tanto, é necessário se ter em mente o conceito básico de que a Densidade Espectral de Potência (*Power Spectral Density (PSD)*),  $S_n(f)$ , pode ser calculada através da transformada de Fourier da função de autocorrelação,  $r_{xx}$ , mostrado em Eq. (3.25).

$$S_n(f) = \frac{1}{T} \int_0^T r_{xx}(t) e^{-j2\pi ft} dt \quad (3.25)$$

O que se pretende com o filtro é para uma sequência aleatória na entrada seja criada uma função de correlação ou uma dada densidade espectral de potência na saída. Esse processo é representado em Fig. 3.6.

Por definição, sabe-se que as amostras não correlacionadas tem uma PSD constante para uma dada largura de banda igual  $|f| < f_s/2$ , em que  $f_s$  é a frequência de amostragem. Enquanto que a variância,  $\sigma_n^2$ , é por definição a área abaixo de  $S_n(f)$ , que é mostrada em Fig. 3.7, que é igual a Eq. (3.26).

$$\sigma_n^2 = \frac{1}{2} N_0 f_s \quad (3.26)$$

É preciso a utilização de resultado básico da teoria de processos estocásticos para se alcançar

uma dada PSD na saída ao se utilizar uma sequência aleatória. Caso o sistema seja linear, a PSD da entrada,  $S_X(f)$ , e da saída,  $S_Y(f)$ , desses processos se relacionam como em Eq. (3.27), em que  $|H(f)|$  é a função de transferência do sistema linear.

$$S_Y(f) = |H(f)|^2 S_X(f) \quad (3.27)$$

Sabe-se que para sequências de ruído independentes a PSD pode ser descrita por uma constante cujo valor é igual a  $K = N_o/2$  Watts/Hz como mostrado em Fig. 3.7. Assim, substituindo a PSD de entrada em Eq. (3.27), chega-se em Eq. (3.28).

$$|H(f)| = \sqrt{S_Y(f)/K} \quad (3.28)$$

Assim, o problema de construção do canal de Jakes se reduz a construção de um filtro cuja função de transferência, que se relaciona através da raiz quadrada da densidade espectral de potência desejada, é descrita em Eq. (3.28). A densidade espectral de potência, que descreve o canal de Jakes, já foi mostrada e está presente em Eq. 3.19.

A construção de filtros, segundo [6], é comumente feita a partir de três tipos de modelo: modelo autorregressivo, *autoregressive* (AR), que possui somente os coeficientes  $a$ ; o modelo de média móvel, *moving average* (MA), que possui apenas os coeficientes  $b$ ; e, por fim, o modelo autorregressivo de média móvel, *autoregressive moving average* (ARMA), que possui os coeficientes  $a$  e  $b$ . Esses coeficientes são oriundos da representação da função de transferência em sua forma digital usada para simulação, que é alcançada por meio da aplicação da transformada Z. A representação da função de transferência em sua forma digital é apresentada em Eq. (3.29).

Os diagramas de representação da construção dos filtros AR, MA e ARMA são mostrados ainda em Fig. 3.8. Ressalta-se ainda que os modelos MA são semelhantes a um modelo FIR, como o que foi usado neste trabalho, pois possuem apenas os coeficientes  $b$  presentes em Eq. (3.29).

$$H[z] = \frac{\sum_{k=0}^K b[k]z^{-k}}{\sum_{l=0}^L a[l]z^{-l}} \quad (3.29)$$

Neste trabalho, o filtro foi implementado por meio de um modelo FIR, ou seja, apresenta apenas coeficientes  $b[n]$ .

### 3.7 Resultados parciais e conclusões

A construção do modelo no MATLAB foi feita baseada no código fornecido no Apêndice A de [5]. Foi analisado o comportamento do canal ao ser inserido um impulso unitário em sua entrada. A resposta ao impulso é mostrada em Fig. 3.9. Além disso, é representado também em Fig. 3.9 a PSD do filtro em dB, verifica-se que o filtro implementado de fato apresenta uma curva semelhante a esperada em formato de  $U$ , que se assemelha a curva mostrada em Fig. 3.3

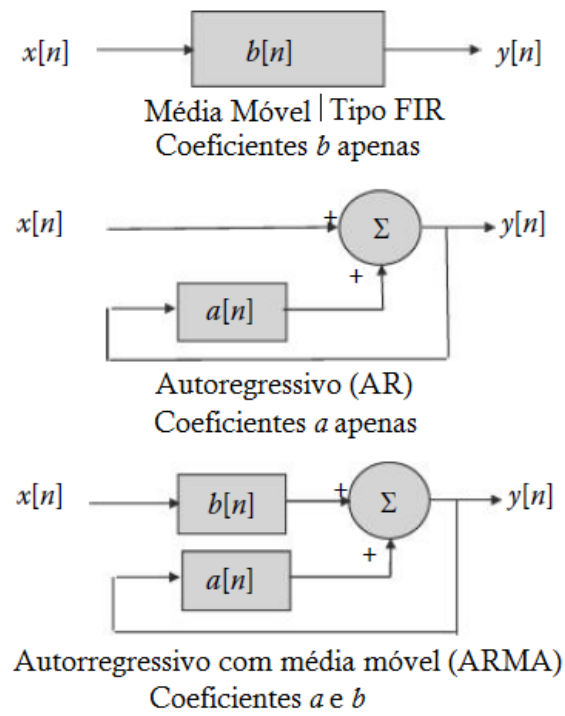


Figura 3.8: Tipos de filtros comumente usados. Adaptado de [6].

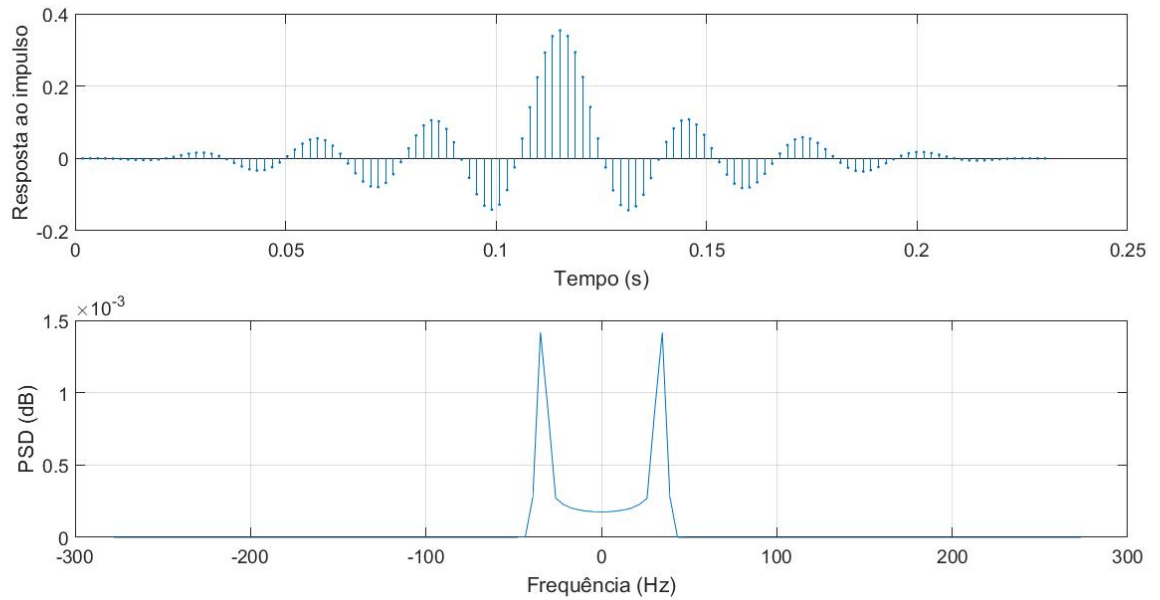


Figura 3.9: Resposta ao Impulso do canal e PSD.

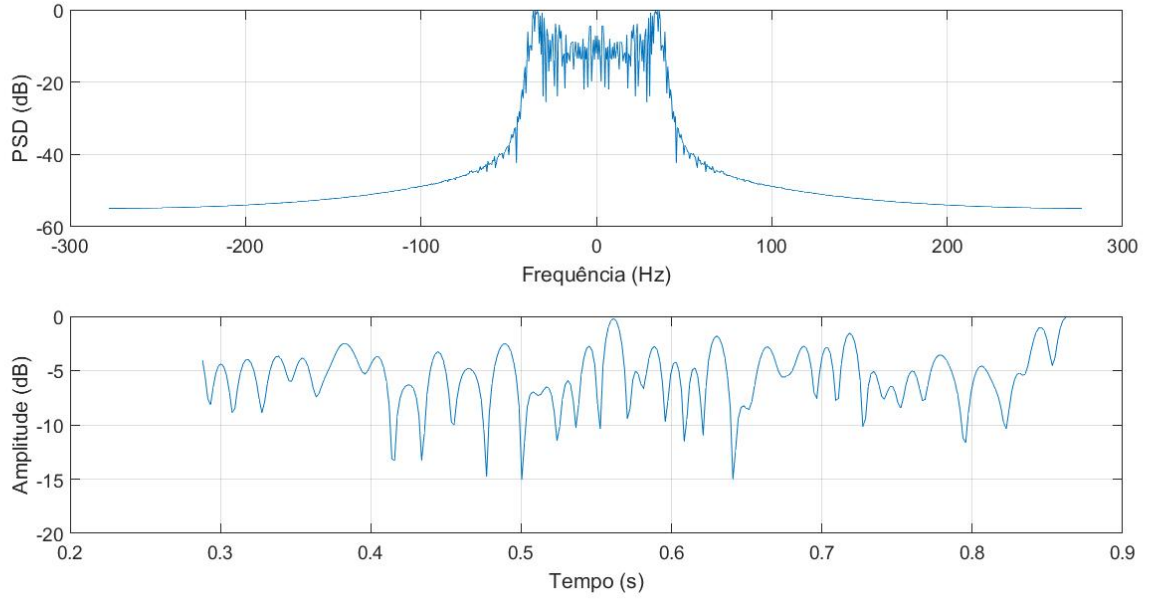


Figura 3.10: PSD para entrada como um ruído e realização temporal do canal.

Ademais, foi verificado também o comportamento do canal quando é inserido um ruído gaussiano real com variância unitária e média zero na entrada do filtro como pode ser visto em Fig. 3.10. Neste caso, foi calculado novamente o PSD na saída do filtro. Verifica-se, novamente, a semelhança da curva com uma letra *U*, que por sua vez, é semelhante ao mostrado em Fig. 3.9. Essa verificação também valida a construção do canal de Jakes por meio de um filtro, em que na entrada desse é inserido um ruído gaussiano não correlacionado para se construir o espectro de potência desejado.

Ainda em relação a Fig. 3.10, há a presença da envoltória do canal na segunda imagem, que como pode se verificado possui um comportamento Rayleigh como era desejado. Com objetivo ainda de se validar essa distribuição da envoltória foi feita a curva que representa a densidade de probabilidade acumulada, que é mostrada em Fig. 3.11

Por fim, foi verificado se o canal de fato foi capaz de gerar a correlação desejada. Para tanto, foi utilizado um ruído gaussiano complexo cujas componentes real e imaginária são independentes de média zero e variância unitária. Em seguida, foi calculada a autocorrelação desse sinal, que espera-se que seja máxima apenas no ponto em que se compara o sinal com ele mesmo sem que haja qualquer deslocamento uma vez que esse foi gerado de maneira aleatório. Isso é de fato o que é observado, como pode ser visto em Fig. 3.12. Em seguida, após a passagem desse sinal pelo processo de filtragem, foi medida novamente a correlação. Nesse caso, espera-se que o comportamento seja próximo a uma função de Bessel de primeira espécie e ordem zero, como descrito em Eq. (3.16), que é observado em Eq. (3.13).

De posse de um modelo de canal, é possível a aplicação de métodos de rastreamento que é o foco desse trabalho para se analisar quais as melhores formas de se estimar um canal de comunicação para que se consiga a maximização dos benefícios trazidos pelo uso das técnicas MIMO massivo.

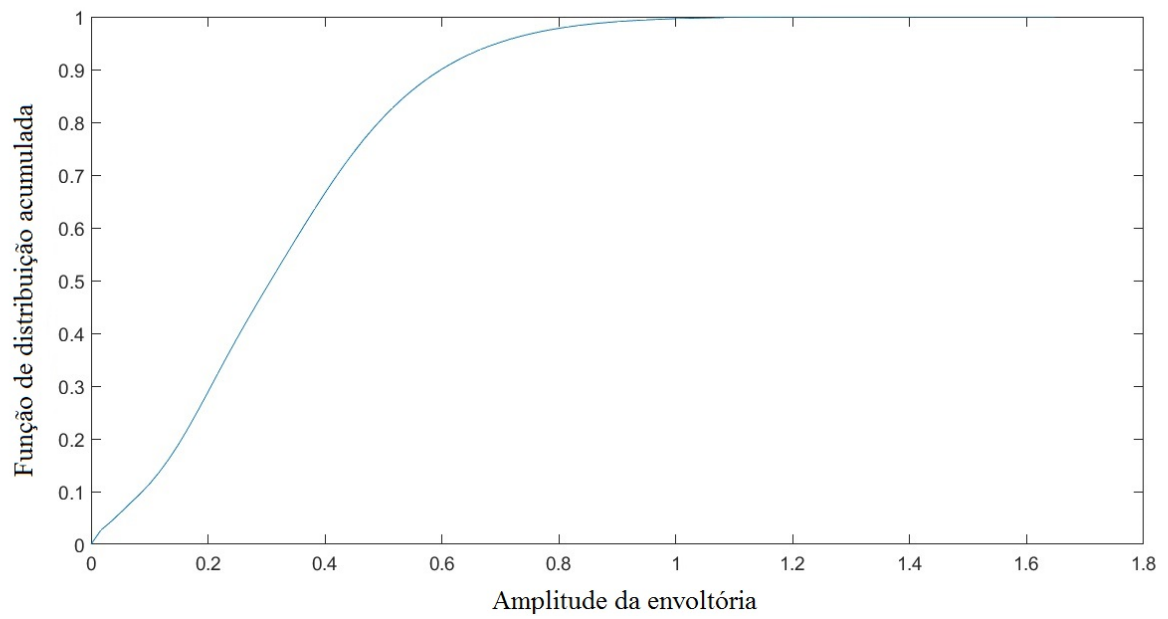


Figura 3.11: Resposta ao Impulso do canal

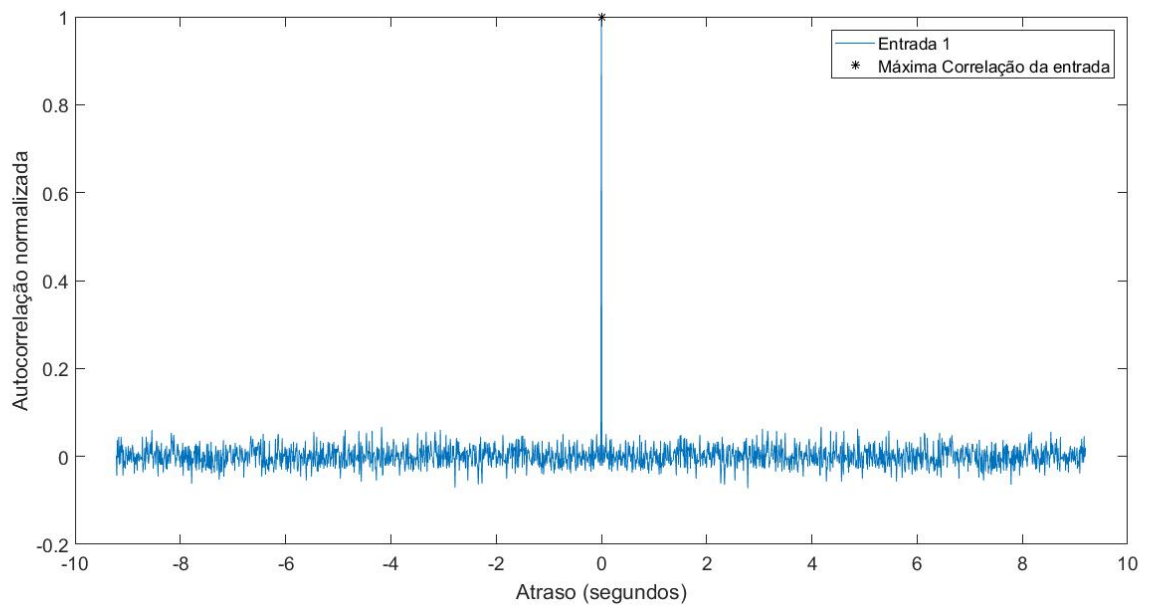


Figura 3.12: Autocorrelação do ruído gaussiano complexo em que as componentes reais e imaginárias são independentes.



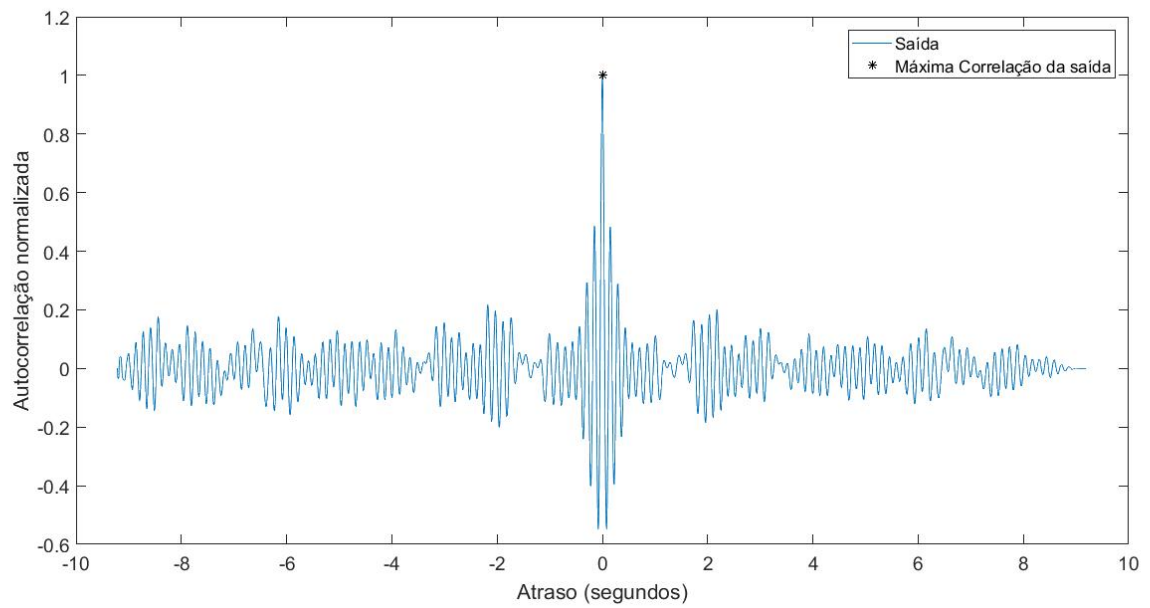


Figura 3.13: Autocorrelação do ruído gaussiano complexo em que as componentes reais e imaginárias são independentes após o processo de filtragem.

Por isso, o capítulo que segue tratará das formas de estimação que serão abordadas nesse trabalho que é o filtro de Kalman e o algoritmo LMS.

## Capítulo 4

# Algoritmo LMS para Rastreo de Canal

*Este capítulo fará uma descrição de um algoritmo de estimação de canal, o algoritmo LMS. Ao fim, serão mostrados os resultados do rastreamento do canal de comunicação descrito no capítulo 3 usando esse método.*

### 4.1 Introdução

Como mostrado anteriormente no capítulo 2, uma correta estimação dos canais em sistemas MIMO é necessária para que os benefícios da utilização de arranjos com centenas de antenas sejam ampliados. Uma vez que o canal que será rastreado já foi gerado segundo descrição do capítulo 3, segue-se então para a descrição dos algoritmos responsáveis pelo rastreamento do canal de comunicação.

O algoritmo LMS remota a década de 60 e tem como objetivo a redução do erro quadrático médio entre a resposta desejada e a obtida por meio do algoritmo [7]. A ideia básica por trás desse modelo é que, a partir de uma superfície que representa o erro de estimação, sejam feitos ajustes dos coeficientes de forma que a derivada da função que representa o erro se oponha ao gradiente da superfície. Isso tem como objetivo, portanto, se opor a direção de maior crescimento da função e se reduzir o erro entre o sinal estimado e o sinal real.

O capítulo será estruturado da seguinte forma: primeiramente, haverá a introdução ao algoritmo do gradiente descendente por meio de um exemplo. Em seguida, haverá o desenvolvimento desse método, que, por sua vez, possibilitará se chegar no algoritmo LMS. Por fim, serão mostrados resultados envolvendo esse método de rastreamento de canal.

### 4.2 Problema para redução do erro quadrático médio

O desenvolvimento do algoritmo LMS será feito a partir do método do gradiente descendente, *Steepest Descent*, desenvolvido com base nas equações de Wiener-Hopf para um processo de filtragem linear ótimo. Para a descrição dessas equações será usado novamente um problema auxiliar.

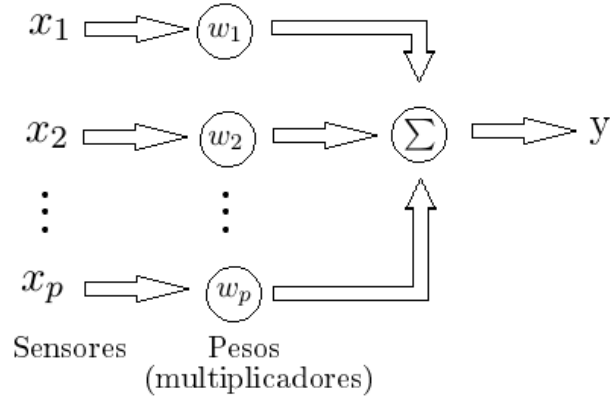


Figura 4.1: Representação espacial do filtro.

A referência usada como base nesta seção é [7].

Neste problema considera-se a existência de  $p$  sensores localizados em diferentes pontos do espaço como mostrado em Fig. 4.1. Cada um dos sensores produz um sinal que são denominados  $x_1, x_2, \dots, x_p$ . Esses são então multiplicados por pesos denominados  $w_1, w_2, \dots, w_p$ , por fim, a saída,  $y$ , é obtida somando-se os sinais multiplicados pelos pesos como mostrado em Eq. (4.1). A resposta desejada é representada por  $d$ , o sinal de erro pode ser definido então como mostrado em Eq. (4.2).

$$y = \sum_{k=1}^p w_k x_k \quad (4.1)$$

$$e = d - y \quad (4.2)$$

A medição do desempenho pode ser feita através de uma função de custos, que é mostrada em Eq. (4.3) e tem como base a média erro quadrático médio. Nesta equação, o termo  $1/2$  foi inserido simplesmente por conveniência.

O problema neste caso é, portanto, determinar os pesos ótimos de forma a minimizar o erro quadrático médio, cuja média é representada pela função de custos,  $J$ . Este problema é denominado na literatura de processamento de sinais como o filtro de Wiener.

$$J = \frac{1}{2} E[e^2] \quad (4.3)$$

### 4.3 Método do Gradiente Descendente

Substituindo-se Eq. (4.1) em Eq. (4.2), tem-se por resultado Eq. (4.3), no qual se fazendo a devida expansão chega-se em Eq. (4.4).

$$J = \frac{1}{2}E[d^2] - E\left[\sum_{k=1}^p w_k x_k d\right] + \frac{1}{2}E\left[\sum_{j=1}^p \sum_{k=1}^p w_j w_k x_j x_k\right] \quad (4.4)$$

Uma vez que a esperança é um operador linear pode-se reescrever Eq. (4.4) como em Eq. (4.5), em que os pesos são considerados constantes e, por isso, retirados do operador esperança.

$$J = \frac{1}{2}E[d^2] - \sum_{k=1}^p w_k E[x_k d] + \frac{1}{2} \sum_{j=1}^p \sum_{k=1}^p w_j w_k E[x_j x_k] \quad (4.5)$$

O primeiro termo presente em Eq. (4.5),  $E[d^2]$ , é o erro quadrático médio da resposta desejada  $d$ , que assumirá uma nova notação mostrada em Eq. (4.6).

$$r_d = E[d^2] \quad (4.6)$$

Já o segundo termo presente em Eq. (4.5),  $E[dx_k]$ , é a função de correlação cruzada entre a resposta desejada  $d$  e o sinal dos sensores  $x_k$ , que assumirá também uma nova notação mostrada em Eq. (4.7).

$$r_{dx}(k) = E[dx_k], \quad \text{para } k = 1, 2, \dots, p \quad (4.7)$$

Por fim, o terceiro termo presente em Eq. (4.5),  $E[x_j x_k]$ , é a função de autocorrelação entre os sinais dos vários sensores, que assumirá também uma nova notação mostrada em Eq. (4.8).

$$r_x(j, k) = E[x_j x_k], \quad \text{para } k, j = 1, 2, \dots, p \quad (4.8)$$

Com essa notação é possível, portanto, reescrever Eq. (4.5) como Eq. (4.9)

$$J = \frac{1}{2}r_d - \sum_{k=1}^p w_k r_{dx}(k) + \frac{1}{2} \sum_{j=1}^p \sum_{k=1}^p w_j w_k r_x(j, k) \quad (4.9)$$

É possível com base em gráfico multidimensional representar a função de custos  $J$  em função dos pesos, que são parâmetros livres. Esse gráfico é denominado superfície de erro do desempenho do filtro. A representação dessa superfície para o caso de apenas um peso é feito em Fig.4.2. O filtro será ótimo quando se alcançar o menor valor de  $J$ . Isso é feito por meio da diferenciação da função de custos com respeito aos pesos  $w_k$ , que também é denominado gradiente da superfície de erro com respeito a um dado peso. Considerando-se que  $\nabla_{w_k} J$  representa o gradiente e é mostrado em Eq. (4.10). O resultado dessa diferenciação é mostrado em Eq. (4.11)

$$\nabla_{w_k} J = \frac{\partial J}{\partial w_k}, \quad \text{para } k = 1, 2, \dots, p \quad (4.10)$$

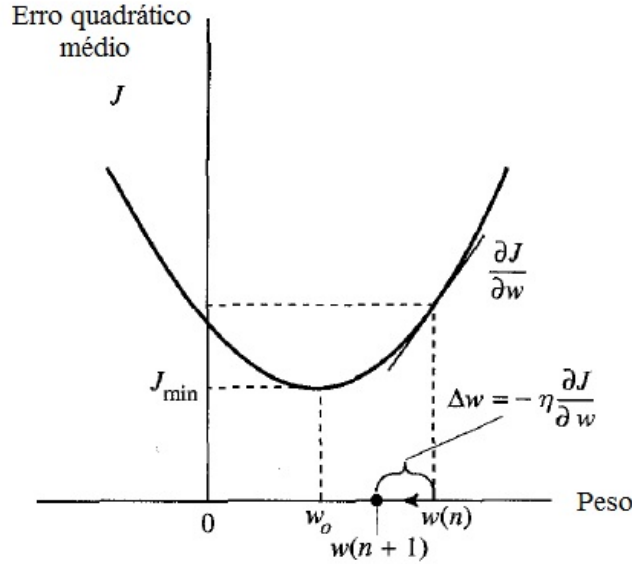


Figura 4.2: Representação da superfície de erro para caso de um peso apenas [7]

$$\nabla_{w_k} J = -r_{dx}(k) + \sum_{j=1}^p w_j r_x(j, k) = 0 \quad \text{para } k = 1, 2, \dots, p \quad (4.11)$$

No valor ótimo para o conjunto de pesos, esses são denominados  $w_{ok}$  e são mostrados em Eq. (4.12), a qual, na verdade, retrata um conjunto de equações denominados *equações de Wiener-Hopf*. Quando um filtro satisfaz este conjunto de equações é denominado filtro de Wiener.

$$\sum_{j=1}^p w_{oj} r_x(j, k) = r_{xd}(k) \quad \text{para } k = 1, 2, \dots, p \quad (4.12)$$

A solução das equações de Wiener-Hopf para o caso do problema descrito anteriormente é basicamente o cálculo da inversa da matriz  $p \times p$ , composta pelos valores da função de autocorrelação  $r_x(j, k)$  para  $j, k = 1, 2, \dots, p$ .

Para se evitar o cálculo da inversa dessa matriz, faz-se uso do método do gradiente descendente, em que os pesos são variantes no tempo e são ajustados a cada iteração de forma a se aproximar progressivamente da solução ótima. Esse método através dos ajustes sucessivos busca sempre o ponto mais baixo na superfície de erro do filtro. O ajuste dos pesos é feito, portanto, de forma a se opor ao vetor gradiente da superfície, que é o vetor que indica a direção de máximo crescimento da função.

Considerando-se que  $w_k(n)$  representa os pesos na iteração  $n$ , a atualização dos valores pode ser feito como em Eq. (4.13), em que  $\Delta w_k(n)$  é descrito como em Eq. (4.14), em que  $\eta$  é denominado parâmetro da taxa de aprendizado. Portanto, a atualização dos pesos do filtro de Wiener é feita com o valor da iteração anterior mais um fator de correção que é proporcional ao negativo do gradiente da superfície de erro com respeito a um dado peso. O método do gradiente descendente

pode ser representado por fim como mostrado em Eq. (4.15)

$$w_k(n+1) = w_k(n) + \Delta w_k(n), \quad \text{para } k = 1, 2, \dots, p \quad (4.13)$$

$$\Delta w_k(n) = -\eta \nabla_{w_k} J(n), \quad \text{para } k = 1, 2, \dots, p \quad (4.14)$$

$$w_k(n+1) = w_k(n) + \eta \left[ r_{dx}(k) - \sum_{j=1}^M w_j(n) r_x(j, k) \right], \quad k = 1, 2, \dots, p \quad (4.15)$$

Uma das questões fundamentais que envolvem esse problema de minimização do erro quadrático médio é a seleção correta do parâmetro da taxa de aprendizado,  $\eta$ , para que o método do gradiente descendente funcione. Além disso, o método do gradiente descendente necessita do conhecimento das funções de correlação espacial,  $r_{dx}(k)$  e  $r_x(j, k)$  que não estão disponíveis em um ambiente desconhecido, por isso, é necessário estimá-las. O algoritmo *least-mean-square* (LMS) fornece formas de se fazer essas estimações.

## 4.4 Algoritmo LMS

O algoritmo LMS baseia-se na estimação instantânea da função de autocorrelação,  $\hat{r}_x(j, k)$ , e da função de correlação cruzada,  $\hat{r}_{xd}(k)$ . Essas são apresentadas em Eq. (4.16) e Eq. (4.17), em que o sinal circunflexo colocado indica que esses valores são estimados.

$$\hat{r}_x(j, k; n) = x_j(n) x_k(n) \quad (4.16)$$

$$\hat{r}_{dx}(k; n) = x_k(n) d(n) \quad (4.17)$$

Substituindo-se Eq. (4.16) e Eq. (4.17) em Eq. (4.15), obtém-se Eq. (4.18). Reorganizando os termos desta última equação obtém-se Eq. (4.19) e Eq. (4.20), em que  $y(n)$  é mostrado, por fim, em Eq. (4.21).

$$\hat{w}_k(n+1) = \hat{w}_k(n) + \eta \left[ x_k(n) d(n) - \sum_{j=1}^p \hat{w}_j(n) x_j(n) x_k(n) \right] \quad (4.18)$$

$$\hat{w}_k(n+1) = \hat{w}_k(n) + \eta \left[ d(n) - \sum_{j=1}^p \hat{w}_j(n) x_j(n) \right] x_k(n) \quad (4.19)$$

$$\hat{w}_k(n+1) = \hat{w}_k(n) + \eta [d(n) - y(n)] x_k(n), \quad \text{para } k = 1, 2, \dots, p \quad (4.20)$$

$$y(n) = \sum_{j=1}^p \hat{w}_j(n) x_j(n) \quad (4.21)$$

Em Eq. (4.21) os pesos são estimados por isso recebem também o sinal circunflexo. Como pode ser visto, o algoritmo LMS pode ser operado em ambiente desconhecido e não estacionário, cujas variâncias são estimadas. E, como uma consequência do não conhecimento do ambiente em que esse opera os pesos iniciais dados a cada um dos sinais são iguais a zero na inicialização do filtro.

Por outro lado, o método do gradiente descendente opera em um ambiente conhecido, cuja matriz de covariância pode ser determinada com precisão e se pode definir ainda condições iniciais para os pesos. E, como consequência do ambiente conhecido, a trajetória seguida dentro da superfície de erro é bem definida e, geralmente, termina na solução ótima  $w_o$ , quando a escolha do parâmetro da taxa de aprendizado,  $\eta$ , é feita de maneira adequada. Ao passo que no algoritmo LMS a trajetória é aleatória e a medida em que o número de iterações aumenta o movimento aleatório se aproxima de um movimento Browniano.

Pode-se ainda citar o fato de a trajetória dentro da superfície de erro ser bem definida no algoritmo do gradiente descendente, concomitantemente, ser aleatória no algoritmo LMS como uma consequência da função de custo que está sendo minimizada. Enquanto no primeiro caso, a função de custo que está sendo minimizada é um erro quadrático médio, no segundo caso, há uma minimização da função de custos instantânea, cuja acurácia do método aumenta com o número de iterações. E, pelo fato de no primeiro caso a média estar sendo analisada, é preciso que haja um armazenamento de dados para que a função de correlação seja calculada, enquanto que isso não é necessário no segundo caso, portanto, há uma redução da necessidade referente a armazenamentos.

Como visto na descrição acima, é necessária também uma análise quanto ao parâmetro da taxa de aprendizado. Quanto menor esse fator, maior será a capacidade do algoritmo de rastrear com um menor erro o sinal desejado. Todavia, o custo para esse aumento na capacidade de rastreamento é uma menor taxa de adaptação, ou seja, caso o sinal rastreado varie rapidamente, um maior número de interações será demandado para o funcionamento correto do método.

Pode-se ainda, com base em Eq. (4.19), reescrever a equação em forma matricial como expresso em Eq. (4.22), em que  $\mathbf{w}(n)$  é descrito em Eq. (4.23) e  $\mathbf{x}(n)$  é descrito em Eq.(4.24).

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \eta [\mathbf{d}(n) - \mathbf{x}^T(n)\hat{\mathbf{w}}(n)] \mathbf{x}(n) \quad (4.22)$$

$$\hat{\mathbf{w}}(n) = [\hat{w}_1(n), \hat{w}_2(n), \dots, \hat{w}_p(n)]^T \quad (4.23)$$

$$\mathbf{x}(n) = [x_1(n), x_2(n), \dots, x_p(n)]^T \quad (4.24)$$

Reorganizando Eq. (4.22), pode-se chegar em Eq. (4.25). Nota-se nesse ponto que o resultado da iteração  $n+1$  depende do resultado na iteração anterior,  $n$ . Denominando-se o termo entre

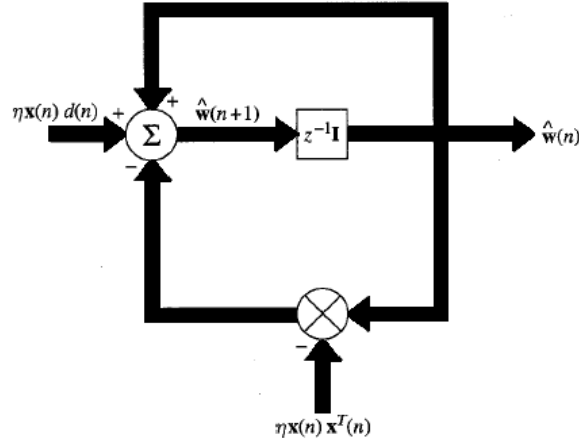


Figura 4.3: Representação em fluxo de sinais do algoritmo LMS [7].

colchetes como operador de atraso unitário,  $\mathbf{z}^{-1}$ . Representando-se então esse algoritmo por meio de um diagrama de fluxo de sinal como em Fig. 4.3. Percebe-se que há a presença de uma malha fechada cuja estabilidade deve ser analisada.

$$\hat{\mathbf{w}}(n+1) = [\mathbf{I} - \eta \mathbf{x}(n) \mathbf{x}^T(n)] \hat{\mathbf{w}}(n) + \eta \mathbf{x}(n) \mathbf{d}(n) \quad (4.25)$$

O problema de convergência do algoritmo LMS pode ser dividido em dois. Por um lado, tem-se a convergência na média do algoritmo, por outro lado, tem-se a convergência no quadrado da média. A primeira é demonstrada no anexo em II.

A convergência na média ocorrerá caso o vetor de pesos estimados  $\hat{\mathbf{w}}(n)$  aproximar-se da solução ótima,  $\mathbf{w}_o$  a medida em que o número de interações aproxima-se de infinito, que é representado em Eq. (4.26). Já a convergência na média quadrada ocorre caso o erro médio quadrático tenda a um valor constante a medida em que o número de iterações aumente, isso é representado em Eq.(4.27)

$$E[\hat{\mathbf{w}}(n)] \rightarrow \mathbf{w}_o, \text{ quando } n \rightarrow \infty \quad (4.26)$$

$$E[e^2(n)] \rightarrow \text{constante}, \text{ quando } n \rightarrow \infty \quad (4.27)$$

## 4.5 Curva de aprendizado

Uma curva em que é apresentado o erro médio quadrático em função do número de iterações é denominado conjunto de média da curva de aprendizado. Esse tipo de curva pode fornecer importantes informações sobre como se dá o comportamento do algoritmo. A curva de aprendizado do algoritmo LMS consiste em uma soma de exponenciais, cuja forma precisa depende da estrutura da matriz de correlação do vetor de entrada [7]. Uma realização apenas, apresenta uma curva



exponencial ruidosa e para que se obtenha uma curva mais suave pode então ser realiza uma média entre várias realizações.

Para um processo que tenha convergência na média quadrática, o erro quadrático médio  $J(n)$  apresenta um valor constante quando o número de interações tende a infinito,  $J(\infty)$ , como visto acima. Isso representa a situação em estado estacionário, que apresenta sempre um erro médio quadrático maior do que o esperado para um filtro de Wiener em um ambiente estacionário,  $J_{min}$ . A diferença entre  $J(\infty)$  e  $J_{min}$  é denominado erro quadrático médio em excesso,  $J_{ex}$ . Essa relação é mostrada em Eq. (4.28). Pode-se ainda definir a razão entre esses erros, apresentada em Eq. (4.29), que é denominada desajuste,  $M$ .

$$J_{ex} = J(\infty) - J_{min} \quad (4.28)$$

$$M = \frac{J_{ex}}{J_{min}} \quad (4.29)$$

Pode-se definir ainda em um filtro LMS o tempo de acomodação, que é o tempo necessário para que o erro fique dentro de uma certa faixa em torno de seu valor quando o número de iterações tende a infinito [17]. Considerando que a exponencial que representa a curva de aprendizado desse sistema tenha a forma apresentada em Eq. (4.30), em que  $J(0)$  seja o erro inicial e  $1/a$  seja a constante de tempo desse sistema, percebe-se que quanto menor o valor da contante de tempo menor será o tempo de assentamento. Isso acarreta, no entanto, um maior erro em regime.

$$J(n) = J(0)e^{-at} \quad (4.30)$$

Ainda segundo [7], o desajuste do algoritmo LMS é diretamente proporcional ao parâmetro da taxa de aprendizado,  $\eta$ , ao passo que o tempo de assentamento é inversamente proporcional. Isso significa que um aumento do parâmetro da taxa de aprendizado gera um menor erro em regime permanente, porém o tempo de acomodação aumenta. Há portanto uma relação de compromisso entre o desajuste do algoritmo LMS e o tempo de acomodação, que é ajustada por meio da taxa de aprendizado. Além disso, como mostrado no anexo II existem limites para que o filtro LMS seja convergente para quaisquer valores iniciais. Apesar disso, a demonstração no anexo II é uma condição necessária para a convergência, no entanto, não é suficiente. É preciso que se verifique a convergência do algoritmo para diferentes valores para a taxa de aprendizagem.

## 4.6 Resultados

O canal de Jakes foi criado no MATLAB tendo como base a descrição feita no capítulo 3. Os parâmetros usados para a geração desse canal foram: a velocidade do movimento relativo entre transmissor e receptor de 3 km/h e a frequência da portadora de 2,5 GHz. Tendo esses parâmetros como base o máximo desvio Doppler usado para a criação do canal de comunicação foi então igual a 6.9444 Hz.

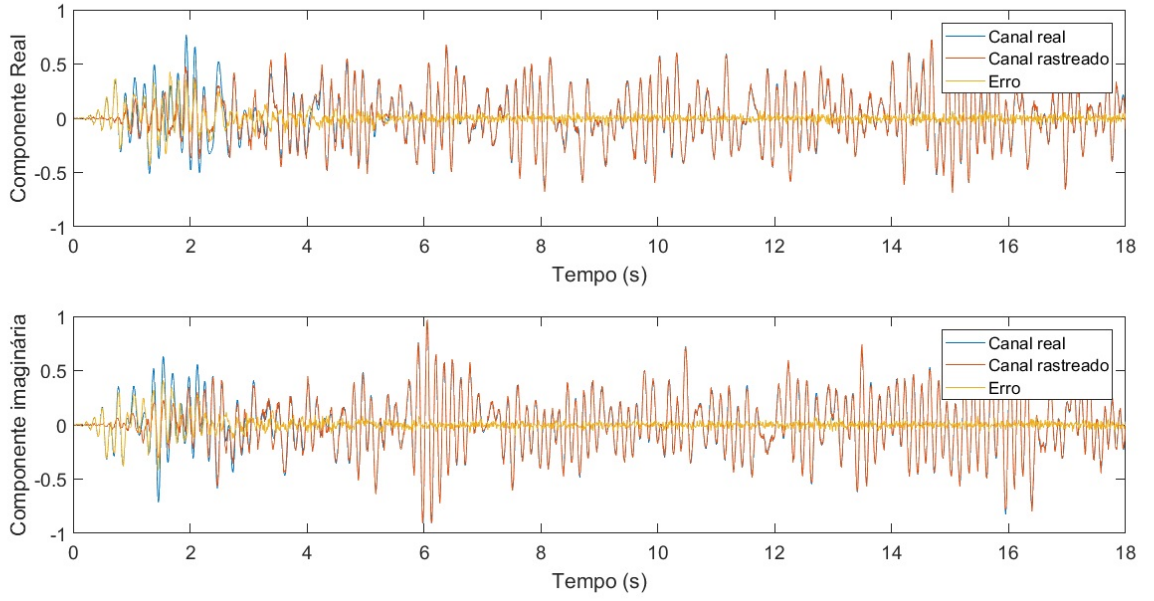


Figura 4.4: Rastreamento do canal de comunicação para taxa de aprendizado = 0,1.

Em seguida, foi necessário escolher de maneira adequada o parâmetro da taxa de aprendizado do algoritmo LMS. Isso porque por um lado, a escolha de um parâmetro com valor muito baixo resultaria em um grande tempo demandado para a estimação do canal de comunicação; por outro lado, a escolha de um parâmetro da taxa de aprendizado muito elevado poderia resultar em um grande erro de estimação do canal de comunicação II.

A necessidade da escolha adequada do parâmetro de aprendizagem do filtro é mostrada com base em Fig. 4.4, Fig. 4.5 e em Fig. 4.6. A relação sinal ruído adotada para se gerar estas figuras foi igual a 20 dB. Nas três figuras, a curva em azul representa o canal de comunicação que deve ser rastreado, a curva vermelha o canal rastreado e a curva em amarelo o erro de estimação.

Em Fig. 4.4, percebe-se que apesar do tempo de estabilização, o filtro rastreia bem o canal de comunicação depois de um certo tempo. Isso porque o valor do parâmetro da taxa de aprendizagem do filtro usado nesse caso foi igual a 0.10 bem próximo ao valor ótimo para a média das realizações do canal com esses parâmetros.

Em Fig. 4.5, há um aumento da taxa de aprendizado para um valor igual a 0,2, que se distancia do valor ótimo. Percebe-se, neste caso, um aumento do erro de estimação do canal. Isso porque um aumento desse parâmetro acima de um determinado valor pode levar a não convergência do filtro como o observado em Fig. 4.6 em que os erros de rastreamento de canal são elevadíssimos para um filtro dessa ordem. A ordem do filtro pode ser vista como o número de sensores do problema inicialmente usado para a descrição do modelo do gradiente descendente. Na construção do algoritmo LMS, a ordem pode ser vista como o número de coeficientes usados para implementá-lo.

Ainda para se mostrar a importância da correta escolha do parâmetro de aprendizado e a relação desse com o erro de estimação foi gerado o gráfico mostrado em Fig. 4.8. Para gerá-lo

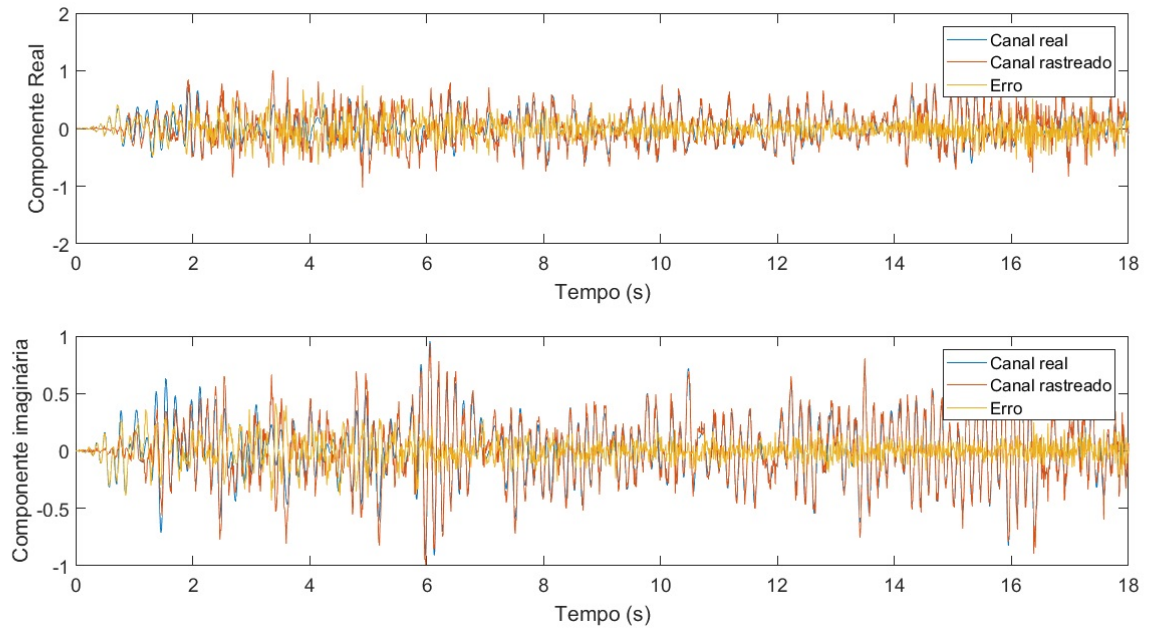


Figura 4.5: Rastreamento do canal de comunicação para taxa de aprendizado = 0,2.

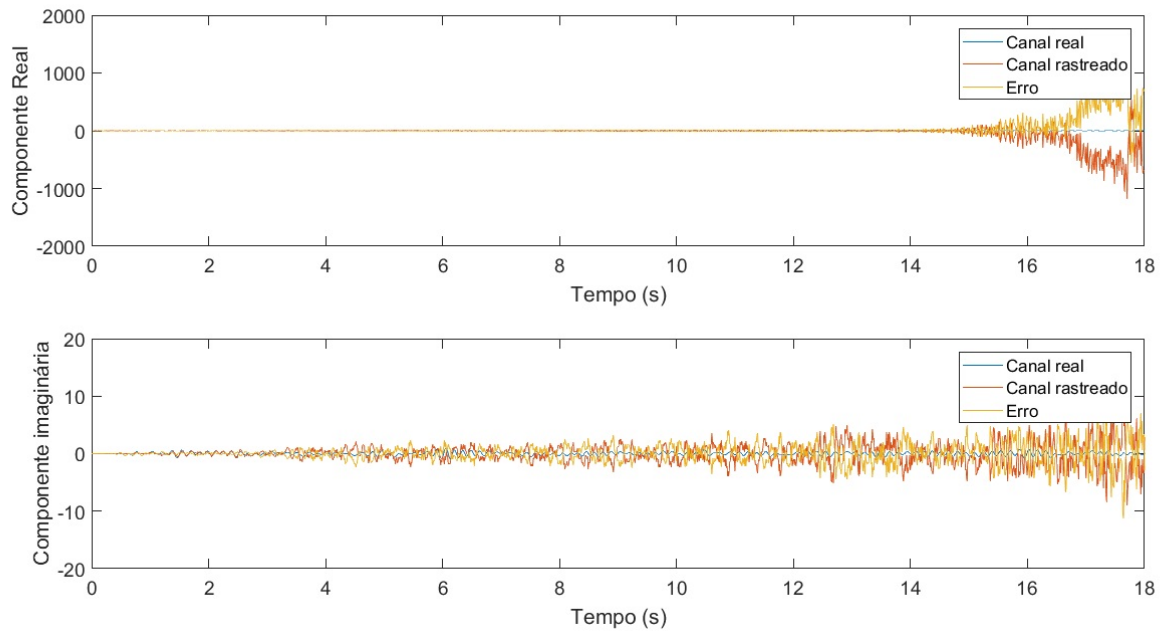


Figura 4.6: Rastreamento do canal de comunicação para taxa de aprendizado = 0,25.

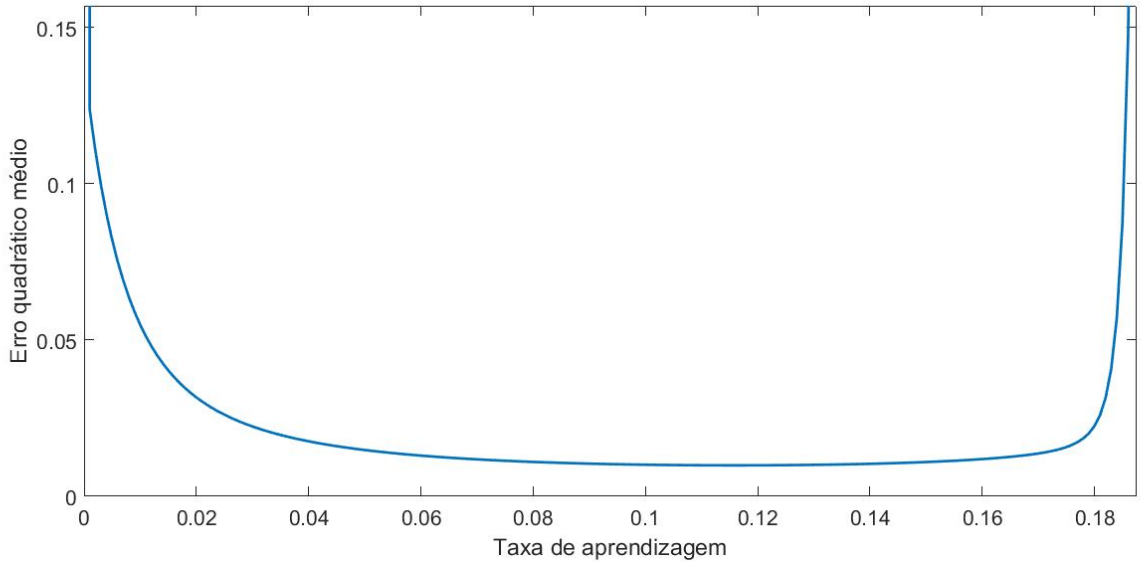


Figura 4.7: Erro quadrático médio pela taxa de aprendizado do algoritmo LMS.

foram feitas 1000 realizações de canal, para cada realização foi calculado o erro médio dada uma taxa de aprendizagem. Por fim, o gráfico foi gerado fazendo uma média dos valores registrados nas 1000 realizações de canal. Percebe que, o parâmetro da taxa de aprendizado muito baixo não permite o correto rastreamento do canal, o que gera erros elevados em relação ao valor ótimo.

Nota-se ainda que valores de aprendizado superiores a 0.18 geram erros de estimação do canal cada vez maiores, que foram suprimidos para que pudesse ser feita também a análise de que existe um valor ótimo para esse parâmetro. O valor ótimo encontrado em uma das simulações com base em 1000 realizações foi de 0.1160 para um filtro de ordem igual a 128. Há ainda uma forte ligação entre a ordem do filtro LMS e o parâmetro da taxa de aprendizado, em que para filtros de ordem mais elevadas é necessário que a taxa de aprendizado seja maior para haver convergência do modelo. Como uma consequência do aumento da ordem, há o aumento do tempo de estabilização do sistema, que é uma característica não desejada.

Foi analisado ainda o rastreamento do canal de comunicação tendo como base diferentes valores de relação sinal ruído, mais especificamente dentro da faixa de 0 a 20 dB. Além disso, foi analisado o impacto da correlação entre canais de comunicação no rastreamento de um canal de comunicação. Essas análises resultaram nas figuras apresentadas em Fig. 4.8 e em Fig. 4.9, em que no primeiro caso foi analisado o impacto da correlação no erro de estimação quando existem apenas 2 antenas transmissoras e no segundo caso foi analisado o caso em que havia 200 antenas transmissoras. Em ambos os casos, percebe-se que há de fato uma melhora na estimação do canal a medida em que a relação sinal ruído aumenta, como era esperado com base em Eq. (2.40).

A figura mostrada em Fig. 4.9 foi feita tendo como objetivo a análise de um sistema MIMO massivo, mais especificamente um sistema MISO, em que há um grande número de antenas transmissoras e apenas uma antena receptora, que visa retratar o caso de um sistema MIMO multi-usuário. Esse sistema seria composto por diversos usuários que tentam fazer a comunicação com

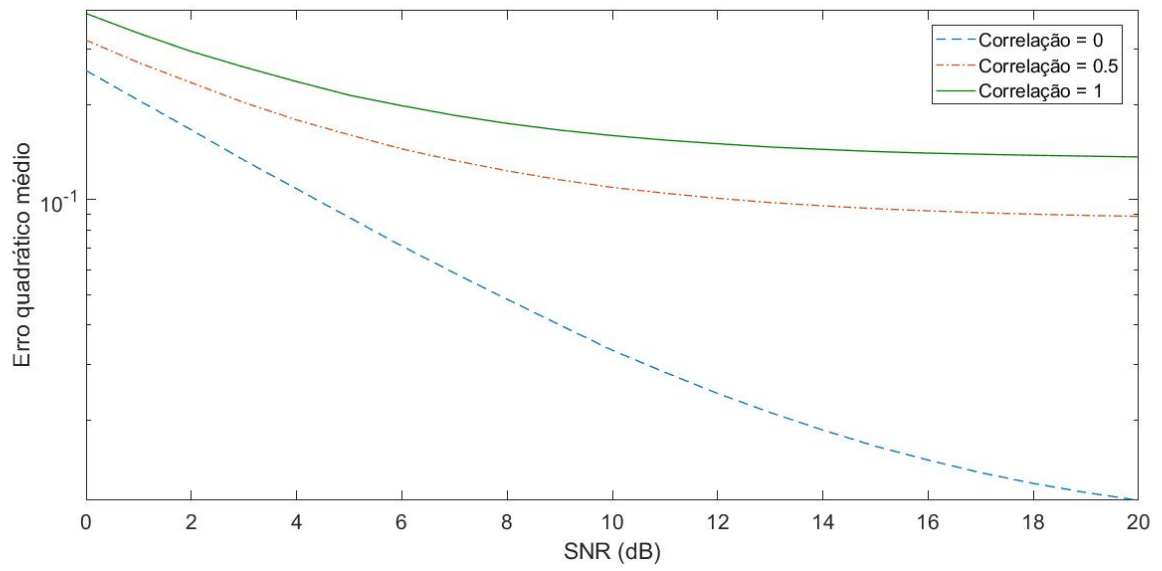


Figura 4.8: Erro quadrático médio pela SNR do algoritmo LMS. Para 100 realizações de canal de canais.  $N_t = 2$ .

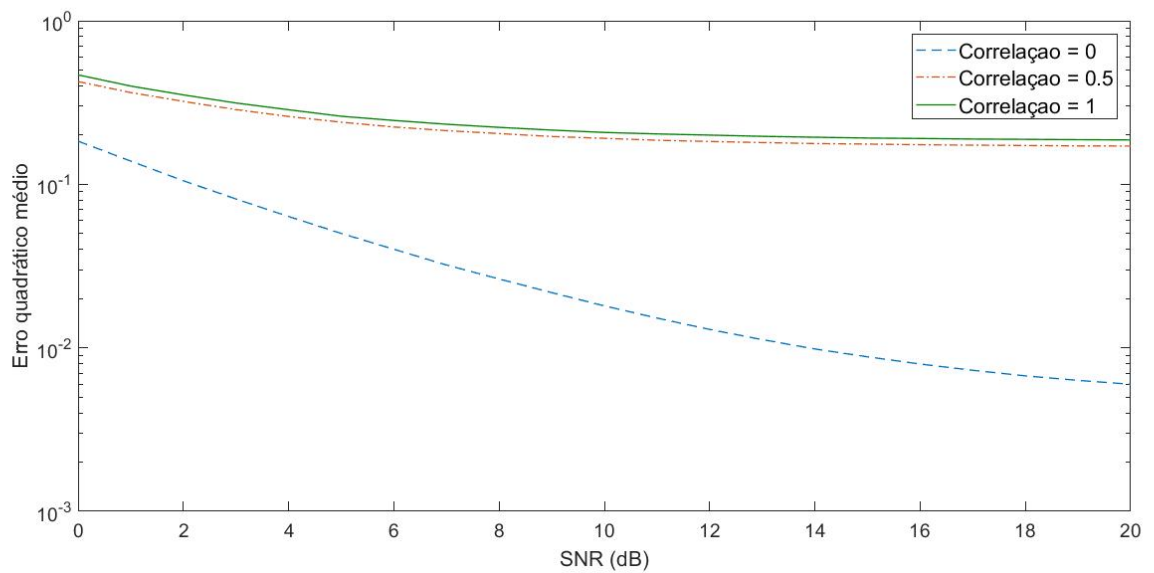


Figura 4.9: Erro quadrático médio pela SNR do algoritmo LMS. Para 100 realizações de canal de canais.  $N_t = 200$ .

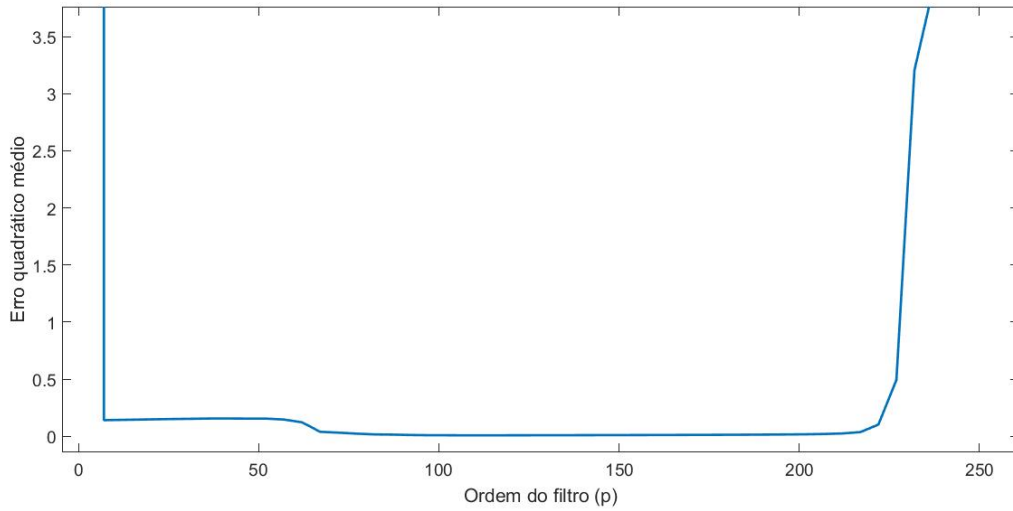


Figura 4.10: Variação da ordem do filtro (p).

a estação base. Percebe-se que de fato a presença de um grande número de canais é responsável por degradar a estimação do canal e, portanto, a capacidade do sistema.

Nota-se, no entanto, ainda com base em Fig. 4.9, que devido ao grande número de canais de comunicação em um sistema MIMO, a correlação entre os canais de comunicação é responsável por rapidamente aumentar o valor do erro quadrático médio do rastreamento do canal. Isso afeta de forma negativa a capacidade do sistema de comunicação de maneira a degradá-la.

O que se espera é que algoritmos de rastreamento sejam otimizados de forma que mesmo quando essa correlação esteja presente entre os canais, o erro no rastreamento não afete de maneira elevada a capacidade do canal de comunicação.

## 4.7 Conclusões parciais

O algoritmo LMS é um filtro de fácil implementação em que não há grandes exigências computacionais, mas para que funcione de maneira correta é necessária a escolha de maneira correta do parâmetro da taxa de aprendizado bem como da ordem do filtro usado para sua implementação.

Analizou-se ainda que a medida em que a relação sinal ruído é aumentada, é possível um melhor rastreamento do canal. Isso significa que seria possível, portanto, uma melhor estimação do canal de comunicação de forma se aumentar a capacidade do canal.

Para um filtro de ordem igual a 128, como o usado neste trabalho, foi feita ainda a escolha do parâmetro de aprendizagem, para que se obtivesse o melhor desempenho possível do filtro. Isso foi feito para que se possa comparar este algoritmo de maneira coerente com o filtro de Kalman, que é apresentado no capítulo seguinte.

## Capítulo 5

# Filtro de Kalman para Rastreo de Canal

*Este capítulo fará uma descrição de um dos algoritmo de estimação de canal, mais especificamente, o filtro de Kalman. Ao fim, serão mostrados os resultados do rastreamento de canal de comunicação descrito no capítulo 3.*

### 5.1 Introdução

O filtro de Kalman, apesar de sua denominação como um filtro, trata-se de um algoritmo computacional, que existe a mais de 50 anos, apesar disso, é ainda um dos algoritmos de fusão de dados mais usados hoje [8]. Isso porque exige pequenos recursos computacionais, apresenta características recursivas elegantes, além de ter o *status* de estimador ótimo de sistema lineares unidimensionais com estatísticas de erro gaussianas. O algoritmo é usado para diversas finalidades entre elas suavização do ruído de informações e permite estimações de parâmetros de interesse [8].

O capítulo será organizado da seguinte maneira: primeiramente, haverá uma descrição intuitiva do filtro de Kalman por meio de um exemplo de rastreamento da posição de um trem. Em seguida, haverá a descrição e demonstração das equações que envolvem esse método. Por fim, serão mostradas simulações envolvendo esse método e os resultados serão comparados com os do algoritmo LMS.

### 5.2 Demonstração intuitiva do filtro de Kalman

Primeiramente, será feito um desenvolvimento intuitivo do filtro de Kalman para melhor compreensão desse algoritmo tendo como base a referência [8]. Por fim, serão demonstradas as equações e a aplicação desse algoritmo ao rastreo de canais de comunicação.

O rastreamento com o uso do filtro de Kalman será inicialmente descrito, de maneira mais intuitiva, para o rastreamento de um trem que se move em uma linha férrea como mostrado em Fig. 5.1. Neste problema, a posição do vagão será estimada pela combinação de duas informações:



a posição fornecida de maneira indireta por um sistema de radio medida por uma antena presente no teto do vagão e a posição estimada com base em valores passados de posição.

No modelo do filtro de Kalman, por ser um filtro recursivo, a posição da locomotiva em um tempo  $t$  envolve a seu valor no tempo  $t - 1$  como mostrado em Eq. (5.1).

$$\mathbf{x}_t = \mathbf{F}_t \mathbf{x}_{t-1} + \mathbf{B}_t \mathbf{u}_t + \boldsymbol{\kappa}_t \quad (5.1)$$

Em Eq. (5.1),  $\mathbf{x}_t$  é o vetor de estados contendo os termos de interesse do sistema,  $\mathbf{F}_t$  é a matriz de transição que faz a relação entre as variáveis do sistema no tempo  $t$  e as variáveis no tempo  $t - 1$ ,  $\mathbf{u}_t$  é o vetor contendo as entradas de controle,  $\mathbf{B}_t$  é a matriz de controle da entrada e, por fim,  $\boldsymbol{\kappa}_t$  é o vetor contendo os termos do ruído do processo para cada parâmetro do vetor de estados.

Já a medição da posição do sistema pode ser analisada por Eq. (5.2). Em que:  $\mathbf{z}_t$  é o vetor de medições;  $\boldsymbol{\Lambda}_t$  é a matriz de transformações que faz a relação do vetor de parâmetros de estados para o domínio das medições e; por fim,  $\mathbf{v}_t$  é o vetor que contém os ruídos de medições.

Tanto o ruído do processo quanto o ruído de medições são ruídos gaussianos brancos com covariância  $\mathbf{Q}_t$  e  $\mathbf{R}_t$  respectivamente. A covariância  $\mathbf{Q}_t$  é definida como em Eq. (5.3) e a covariância  $\mathbf{R}_t$  é definida em Eq. (5.4). Além disso, esses são não correlacionados entre si, como mostrado em Eq. (5.5).

$$\mathbf{z}_t = \boldsymbol{\Lambda}_t \mathbf{x}_t + \mathbf{v}_t \quad (5.2)$$

$$\boldsymbol{\Lambda}_t$$

$$\mathbf{Q}_t = E [\boldsymbol{\kappa}_t \boldsymbol{\kappa}_t^T] \quad (5.3)$$

$$\mathbf{R}_t = E [\mathbf{v}_t \mathbf{v}_t^T] \quad (5.4)$$

$$E [\mathbf{w}_k \mathbf{v}_i^T] = 0, \text{ para todos } k \text{ e } i \quad (5.5)$$

A estimativa da posição é feita pelo sistema de rádio e possui um certo erro associado, que é o erro de medição que é representado por  $\mathbf{v}_t$ . Este sistema de rádio é capaz também de medir a velocidade da locomotiva, que é essencial para o processo de estimação da posição. Essas variáveis de interesse estão presentes no vetor de estados,  $\mathbf{x}_t$ . Isso é mostrado em Eq. (5.6), em que  $x_t$  representa a posição do trem e  $\dot{x}_t$  representa a velocidade do trem, cujo sinal acima de  $x_t$  representa a derivada da posição em relação ao tempo.

Nesse sistema, as entradas podem modificar a posição e a velocidade são a aceleração positiva ou negativa, frenagem. Essas podem ser expressas como sendo iguais as forças aplicadas no trem sobre a massa do trem,  $m$ , como mostrado em Eq. (5.7).

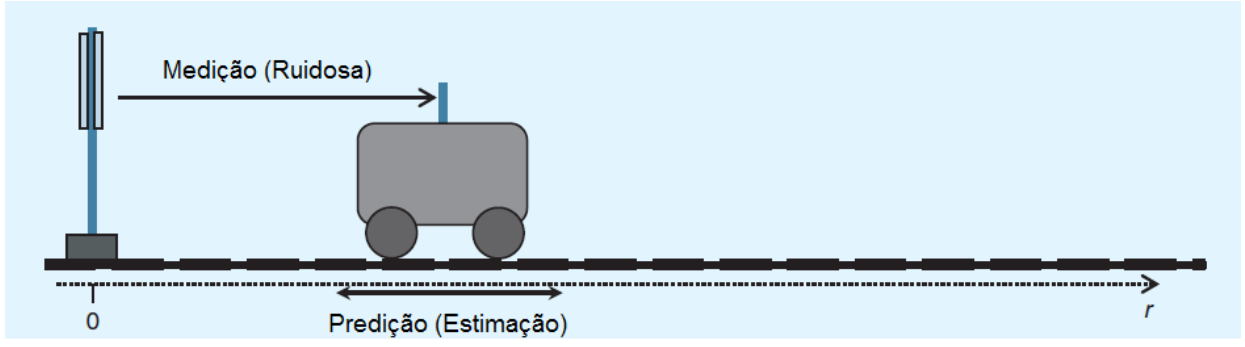


Figura 5.1: Representação do problema. Adaptado de [8].

$$\mathbf{x}_t = \begin{bmatrix} x_t \\ \dot{x}_t \end{bmatrix} \quad (5.6)$$

$$\mathbf{u}_t = \frac{f_t}{m} \quad (5.7)$$

A descrição matemática da posição do trem de acordo com a velocidade e a aceleração pode ser feita pelas equações do movimento uniformemente variado descritas em Eq. (5.8) e Eq. (5.9), que podem ser descritas em sua forma matricial como em Eq. (5.10).

$$x_t = x_{t-1} + \dot{x}_t \times \Delta t + \frac{f_t(\Delta t)^2}{(2m)} \quad (5.8)$$

$$\dot{x}_t = \dot{x}_{t-1} + \frac{f_t(\Delta t)}{(m)} \quad (5.9)$$

$$\begin{bmatrix} x_t \\ \dot{x}_t \end{bmatrix} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ \dot{x}_{t-1} \end{bmatrix} + \begin{bmatrix} \frac{(\Delta t)^2}{2} \\ \Delta t \end{bmatrix} \frac{f_t}{m} \quad (5.10)$$

Comparando-se Eq. (5.10) com Eq. (5.1), percebe-se que  $\mathbf{F}_t$  é descrito como em Eq. (5.11) e  $\mathbf{B}_t$  como em Eq. (5.12).

$$\mathbf{F}_t = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \quad (5.11)$$

$$\mathbf{B}_t = \begin{bmatrix} \frac{(\Delta t)^2}{2} \\ \Delta t \end{bmatrix} \quad (5.12)$$

Além disso, considerando-se que o trem partiu da origem, sabe-se com razoável precisão a posição do trem em seu estado inicial ( $t = 0$ ), que segue uma função distribuição de probabilidade mostrada em Fig. 5.2.

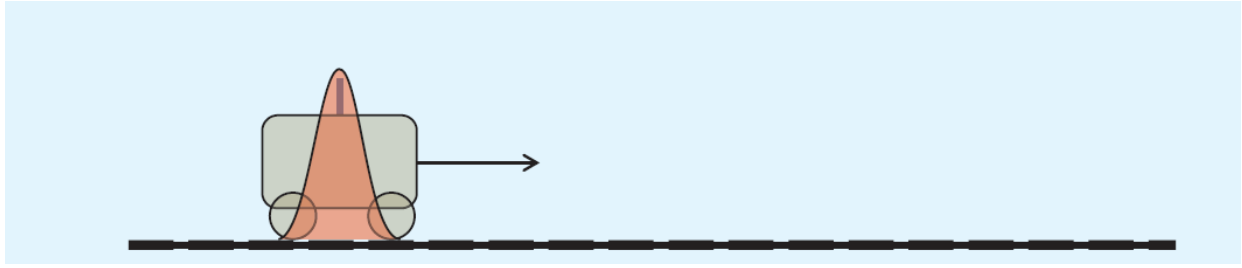


Figura 5.2: Informações iniciais sobre o sistema no tempo  $t = 0$  [8].

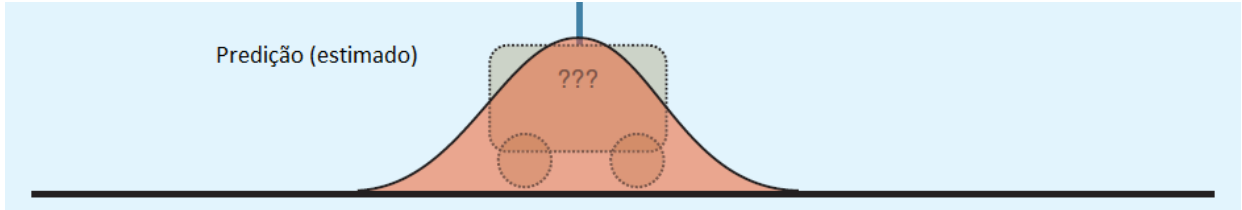


Figura 5.3: Informações iniciais sobre o sistema no tempo  $t = 1$  [8].

A partir da direção e da velocidade iniciais do trem em  $t = 0$ , e nas informações como a máxima aceleração possível do trem, é possível se fazer uma predição da posição em um tempo  $t = 1$ . E, com a posição e velocidade medidos de forma indireta, pode-se fazer a melhor estimativa para a posição do trem que será, portanto, uma combinação desses dois casos, predição e medição. A representação do sistema em  $t = 1$  é mostrada em Fig. 5.3, cujo aumento da variância representa uma redução da acurácia na estimação da posição do trem.

O filtro de Kalman é usado para permitir uma estimação da posição,  $\hat{x}_t$ , combinando a modelagem matemática do sistema, ruídos de medições e funções lineares dos parâmetros. A posição estimada nesse algoritmo não é um valor discreto, mas uma função densidade de probabilidade. Considera-se neste trabalho que o ruído de processo e o ruído de medição seguem distribuições gaussianas. E, para se descrever esse tipo de funções, é necessário se conhecer a matriz de covariância,  $\mathbf{P}_t$ , dos elementos presentes no vetor de estados. Essa covariância, no problema acima descrito, permite que seja determinada a região com maior probabilidade de se encontrar o trem, além disso, permite a determinação dos limites possíveis para a velocidade no tempo  $t = 1$ .

A estimativa da posição, dessa forma, é descrita como uma densidade de probabilidade gaussiana resultado do produto entre as medições e as predições passadas que são também funções de densidade com distribuição gaussiana. Essa característica permite a combinação de um grande número de funções gaussianas sem aumentar a complexidade ou número de termos, como mostrado em Fig. 5.4. Isso confere ao filtro de Kalman uma característica de filtro de características recursivas elegantes [8].

Para o caso de um problema cujo sistema é linear unidimensional cujos erros de medição possuem distribuição Gaussiana com média zero foi mostrado que o filtro de Kalman comporta-se como um estimador ótimo [18].

Os termos inseridos em uma função devem estar em um mesmo domínio, muitas vezes, no

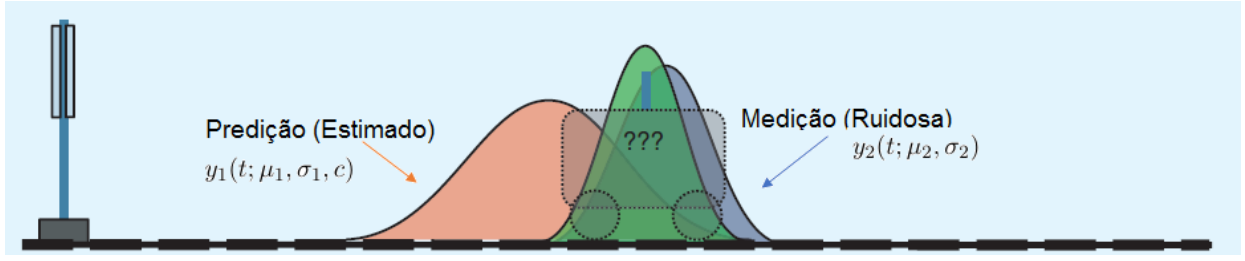


Figura 5.4: Determinação da posição do trem com base em medições ruidosas sobre a posição e predições passadas. Adaptado de [8].

entanto, as etapas de predições e medições estão em domínios diferentes. No caso do problema de estimação da localização do trem, a posição é medida por uma distância, já a localização no sistema de rádio, é feita a partir do tempo que o sinal demorou para percorrer o trajeto entre o trem e a estação base. Assim,  $\mathbf{\Lambda}_t$ , a matriz de transformação, é necessária para mapear as unidades em um mesmo domínio, que no caso do problema da estimação da localização é um distância.

A função distribuição de probabilidade da etapa de predição, é representada em vermelho em Fig. 5.4 e por ser uma distribuição gaussiana pode ser apresentada como em Eq. (5.13). Nesta equação,  $\mu_1$  é a média e  $\sigma_1$  é o desvio padrão. A presença do fator  $c$ , velocidade da luz, é uma consequência da mudança do domínio de metros para tempo para que houvesse uma compatibilidade de unidades entre o valor estimado da posição e o valor.

$$y_1(t; \mu_1, \sigma_1, c) = \frac{1}{\sqrt{2\pi \left(\frac{\sigma_1}{c}\right)^2}} e^{-\frac{\left(t - \frac{\mu_1}{c}\right)^2}{2 \left(\frac{\sigma_1}{c}\right)^2}} \quad (5.13)$$

Ademais, a função distribuição de probabilidade da medição, que também segue uma distribuição gaussiana é representada pela curva em azul em Fig. 5.4 e é dada por Eq. (5.14).

$$y_2(t; \mu_2, \sigma_2) = \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(t - \mu_2)^2}{2\sigma_2^2}} \quad (5.14)$$

O produto entre Eq. (5.13) e Eq. (5.14) permite se chegar em uma nova função densidade de probabilidade que é descrita por Eq. (5.15), em que  $\mu_{final}$  e  $\sigma_{final}$  são relacionados com  $\mu_1$ ,  $\sigma_1$ ,  $\mu_2$  e  $\sigma_2$  como mostrado respectivamente em Eq. (5.16) e Eq. (5.17).

$$y_{final}(t; \mu_{final}, \sigma_{final}) = \frac{1}{\sqrt{2\pi\sigma_{final}^2}} e^{-\frac{(t - \mu_{final})^2}{2\sigma_{final}^2}} \quad (5.15)$$

$$\mu_{final} = \mu_1 + \left[ \frac{\frac{\sigma_1^2}{c}}{\left(\frac{\sigma_1}{c}\right)^2 + \sigma_2^2} \right] \left( \mu_2 - \frac{\mu_1}{c} \right) \quad (5.16)$$

$$\sigma_{2final}^2 = \sigma_1^2 - \left[ \frac{\frac{\sigma_1^2}{c}}{\left(\frac{\sigma_1}{c}\right)^2 + \sigma_2^2} \right] \frac{\sigma_1^2}{c} \quad (5.17)$$

Com as considerações presentes em Eq. (5.18) e Eq. 5.19, pode-se reescrever Eq. (5.16) e Eq. (5.17) como presente em Eq. (5.20) e Eq. (5.21). Estas equações são denominadas as equações do estágio de medições, que possuem a formulação básica mostrada em Eq. (5.22) e Eq. (5.23).

$$H = \frac{1}{c} \quad (5.18)$$

$$K = \frac{(H\sigma_1^2)}{H^2\sigma_1^2 + \sigma_2^2} \quad (5.19)$$

$$\mu_{final} = \mu_1 + \left( \frac{H\sigma_1^2}{H\sigma_1^2 + \sigma_2^2} \right) (\mu_2 - H\mu_1) \quad (5.20)$$

$$\sigma_{final}^2 = \sigma_1^2 - \left( \frac{H\sigma_1^2}{H\sigma_1^2 + \sigma_2^2} \right) H\sigma_1^2 \quad (5.21)$$

$$\hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t(\mathbf{z}_t - \mathbf{\Lambda}_t\hat{\mathbf{x}}_{t|t-1}) \quad (5.22)$$

$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{K}_t\mathbf{\Lambda}_t\mathbf{P}_{t|t-1} \quad (5.23)$$

### 5.3 Descrição matemática do filtro de Kalman

O filtro de Kalman pode ser visto, portanto, como composto por dois estágios básicos um para atualização das predições e outro para atualizações das medições. O primeiro estágio é o responsável por estimar com base em posições anteriores qual será a nova posição do sistema, descrito por Eq. (5.24) e Eq. (5.25), cujas demonstrações são mostradas no anexo em I.1. Essas demonstrações se baseiam na minimização do erro quadrático médio em que o erro é descrito por Eq. (5.26). A minimização do erro quadrático médio pode ser vista como a minimização da matriz de covariância, que é mostrada em Eq. (5.27), já que a covariância é calculada baseada no erro. Na notação, o índice  $t|t-1$  representa estimações em um tempo  $t$  com base em estimações anteriores de  $t-1$ .

$$\hat{\mathbf{x}}_{t|t-1} = \mathbf{F}_t \hat{\mathbf{x}}_{t-1|t-1} + \mathbf{B}_t \mathbf{u}_t \quad (5.24)$$

$$\hat{\mathbf{P}}_{t|t-1} = \mathbf{F}_t \mathbf{x}_{t-1|t-1} \mathbf{F}_t^T + \mathbf{Q}_t \quad (5.25)$$

$$\mathbf{e}_{t|t-1} = x_t - \hat{x}_{t|t-1} \quad (5.26)$$

$$\mathbf{P}_{t|t-1} = E \left[ (\mathbf{e}_{t|t-1})(\mathbf{e}_{t|t-1})^T \right] \quad (5.27)$$

Já o segundo estágio, de medições, é responsável por combinar a estimação do estágio de predição com a medição feita de forma indireta e é descrita por Eq. (5.28) e por Eq. (5.29). A demonstração de Eq. (5.29) está presente no anexo em I.1.

$$\hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t(\mathbf{z}_t - \mathbf{\Lambda}_t \hat{\mathbf{x}}_{t|t-1}) \quad (5.28)$$

$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{K}_t \mathbf{\Lambda}_t \mathbf{P}_{t|t-1} \quad (5.29)$$

Em Eq. (5.28),  $\mathbf{z}_t$  é o vetor de medições de um tempo  $t$  e é mostrado em Eq. (5.30);  $\mathbf{K}_t$  é um fator que faz a multiplicação da diferença entre o valor estimado e o valor anterior para que se aproxime após a etapa de medições o máximo possível do valor real. Esse ganho é denominado ganho de Kalman quando se minimiza os mínimos quadrados, a demonstração deste é mostrado no anexo I.1, e é mostrado em Eq. (5.31). Por fim, nesta etapa a covariância é atualizada pela expressão em Eq. 5.32.

$$\mathbf{z}_t = \mathbf{\Lambda}_t \mathbf{x}_t + \mathbf{V}_t \quad (5.30)$$

$$\mathbf{K}_t = \mathbf{P}_{t|t-1} \mathbf{\Lambda}_t^T (\mathbf{\Lambda}_t \mathbf{P}_{t|t-1} \mathbf{\Lambda}_t^T + \mathbf{R}_t)^{-1} \quad (5.31)$$

Com base nesse conjunto de equações é possível a utilização do filtro de Kalman, cujo esquemático que mostra o funcionamento do filtro é mostrado em Fig. 5.5.

$$\mathbf{P}_{t|t-1} = \mathbf{F}_t \mathbf{P}_{t|t} \mathbf{F}_t^T + \mathbf{Q}_t \quad (5.32)$$

## 5.4 Filtro de Kalman aplicado a estimação de canais

O canal de comunicação neste trabalho foi descrito por um modelo de primeira ordem [3] mostrado em Eq. (5.33) em que  $\mathbf{x}_t = \text{vec}(H_t)$ , cujas dimensões são de  $N_t N_r \times 1$  para um  $k$ -ésimo

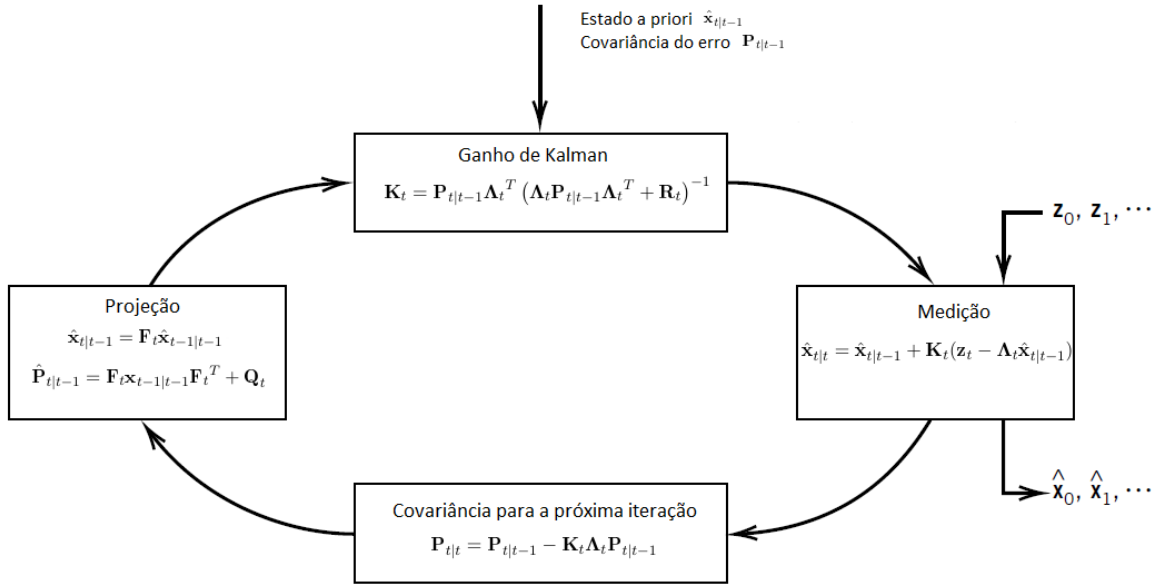


Figura 5.5: Ciclo do filtro de Kalman.

bloco;  $U_t$  é denominado vetor de inovação, que apresenta distribuição Gaussiana complexa de média zero no e variância igual a  $R_h$ , que por sua vez é mostrado em Eq. (5.34).

A notação que permite a representação dessa distribuição é mostrada em Eq. (5.35). E, como as componentes que chegam ao receptor são identicamente distribuídas, apenas os elementos na sua diagonal principal da matriz  $U_t$  são diferentes de zero. Supõe-se ainda que  $h_o$  apresenta distribuição semelhante a de  $U_t$ .

$$x_{t|-1} = A x_{t-1|t-1} - \sqrt{1 - \alpha^2} U_t \quad (5.33)$$

$$R_h = E [x_t x_t^T] \quad (5.34)$$

$$U_t \sim N(0, R_h) \quad (5.35)$$

A matriz  $A$ , para o caso em que não há correlação entre os canais de comunicação, é uma matriz diagonal, cujos termos são iguais a  $\alpha$ . O parâmetro  $\alpha$  pode assumir valores dentro do intervalo  $(0, 1]$ , nesse caso  $\alpha = J_0(2\pi f_D T_s)$ , em que  $J_0$  é a função de Bessel do primeiro tipo e ordem zero,  $T_s$  é o intervalo de símbolo transmitido e  $f_D$  é o máximo desvio Doppler.

Nota-se que o coeficiente de desvanecimento temporal,  $\alpha$ , é determinado pelo movimento relativo entre transmissor e receptor, dessa forma, por um lado, quando não houver movimento relativo, o canal permanecerá invariante no tempo ( $\alpha = 1$ ); por outro lado, quanto maior for o movimento relativo, menor será valor do coeficiente de desvanecimento temporal, tendo o valor zero como limite [19].

O sinal recebido é apresentado em Eq. (5.36), em que  $\mathbf{Y}_t$  é a matriz de sinais recebidos de dimensões  $N_r \times T$ ,  $T$  que por sua vez representa o comprimento dos símbolos transmitidos. Nesse trabalho  $T$  será igual ao número de antenas transmissoras. Ademais,  $\mathbf{H}_t$  é a matriz de ganhos complexos, cujas dimensões são  $N_r \times N_t$ . Por fim,  $\mathbf{N}_t$  é a matriz que contém amostras de ruído gaussiano complexo independente e identicamente distribuídos.

$$\mathbf{Y}_t = \mathbf{H}_t \mathbf{X}_t + \mathbf{N}_t \quad (5.36)$$

O sinal recebido é apresentado em Eq. (5.36) pode ser rescrito como em Eq. (5.37), em que  $\mathbf{S}_k^H = \mathbf{X}_t \otimes \mathbf{I}_{N_r}$  é uma matriz de dimensões  $N_r N_t \times N_r N_t$  é composta por pilotos, usados para estimar o canal;  $\mathbf{y}_t = \text{vec}(\mathbf{Y}_t)$  e  $\mathbf{n}_t = \text{vec}(\mathbf{N}_t)$ . A mudança na ordem de apresentação dos termos entre Eq. (5.36) e Eq. (5.37) foi feita apenas para melhor visualização da relação entre esse modelo e o descrito anteriormente para o filtro de Kalman.

Nesse trabalho, a sequência de pilotos foi escolhida como sendo uma coluna de uma matriz identidade, pois seria a forma mais simples e eficaz de se obter pilotos ortogonais.

$$\mathbf{y}_t = \mathbf{S}_t^H \mathbf{h}_t + \mathbf{n}_t \quad (5.37)$$

Fazendo-se uma relação do modelo de canal descrito nessa seção com o do filtro de Kalman, percebe-se que a matriz  $\mathbf{A}$  de Eq. (5.33) equivale a matriz  $\mathbf{F}_t$  de Eq. (5.24), enquanto que o termo  $\sqrt{1 - \alpha^2}$  representa  $\mathbf{B}_t$ . É possível fazer de maneira análoga uma comparação entre Eq. (5.37) e Eq. 5.30.

## 5.5 Resultados

Para que o processo de rastreamento com o uso do filtro de Kalman pudesse ser realizado, foi preciso criar uma planta do processo. Essa planta do processo seria então a representação no espaço de estado do canal de comunicação. Esse fato é de extrema relevância para analisar se o erro de rastreamento do canal é uma consequência do modelo de canal usado na planta ou uma consequência do rastreamento por meio do uso do filtro de Kalman.

A partir da construção da planta do processo, foi possível a introdução dessa no filtro de Kalman uma vez que este também possui suas bases definidas na representação no espaço de estados. A representação do algoritmo usado é apresentado em Fig. 5.6. Neste trabalho, no entanto, o ruído do processo foi considerado como sendo nulo.

Com base no diagrama mostrado em Fig. 5.6, foi possível o cálculo de dois erros quadráticos médios, em que um deles é a diferença da saída do filtro de Kalman,  $y_e$ , e a saída da planta sem a presença do ruído do sensor,  $y$ . Já o segundo erro quadrático médio pode ser calculado entre o processo introduzido na planta,  $u$ , que é o canal de fato deve a rastreado e a saída do filtro de Kalman,  $y_e$ .



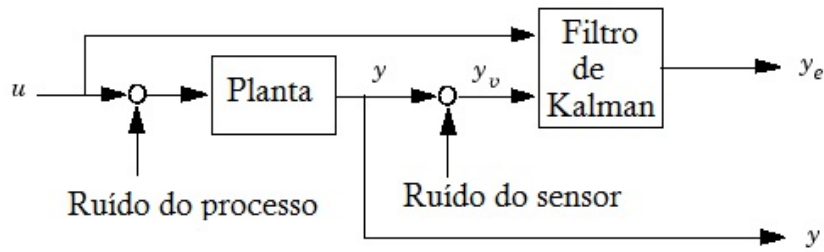


Figura 5.6: Diagrama do rastreamento do canal de comunicação com o uso do filtro de Kalman.

Em Fig. 5.7, são mostrados os rastreamentos da componente real e imaginária do ganho do canal em que é apresentado o canal de comunicação que deve ser rastreado em verde, o canal que sai da planta é a curva tracejada em azul, por fim, o canal rastreado pelo filtro de Kalman é apresentado em vermelho. Percebe-se que, desde o início do processo, o filtro de Kalman é capaz de acompanhar o processo. Mas a passagem do canal de comunicação na planta é responsável por um atraso no canal estimado em relação ao canal que de fato é apresentado. Isso acarreta em um maior erro quadrático médio entre o canal real e o canal rastreado pelo filtro. Nota-se ainda que o filtro de Kalman apresenta um tempo de acomodação muito menor do que o encontrado no algoritmo LMS.

Foi feita em Fig. 5.8, uma simulação do comportamento do erro quadrático médio do filtro quando a relação sinal ruído é alterada com base em 1000 realizações do canal. Esse erro foi calculado entre a saída do filtro de Kalman,  $y_e$ , e a saída da planta,  $y$ . Percebe-se que o erro observado nesse caso foi bem menor do que o observado no rastreamento com o algoritmo LMS.

Tudo isso vai ao encontro do que é encontrado na literatura em que o filtro de Kalman é apresentado como um algoritmo com excelente desempenho para estimação de sistemas. No entanto, como observado em Fig. 5.9, quando é medido o erro entre o canal estimado pelo filtro de Kalman,  $y_e$ , e o canal real,  $u$ , que de fato deve ser rastreado, percebe-se que o erro é mais elevado uma vez que há a presença também do erro introduzido pelo modelo de sistema presente na planta.

Um dos elementos presente no modelo do espaço de estados na planta que afeta o desempenho da saída da planta em acompanhar o canal real é o coeficiente de desvanecimento temporal,  $\alpha$ . Como forma de se mensurar o impacto desse parâmetro, foi feita uma curva de erro quadrático médio entre o canal real e de saída da planta, que é apresentada em Fig. 5.10. Percebe-se que, quanto menor o valor do coeficiente de desvanecimento, menor o erro registrado. Isso pode ser explicado uma vez que quanto menor esse parâmetro, maior é a taxa de variação do canal de comunicação e, portanto, menor é a dependência da estimação atual do valor anterior.

Como o algoritmo usado para a implementação do método LMS não previa esse atraso em relação ao sinal que deveria ser estimado na implementação do filtro de Kalman existe ainda a presença desse erro do modelo no espaço de estado, isso pode gerar resultados em que o filtro de Kalman não apresente seu melhor desempenho. Em Fig. 5.11, foi feita uma correção temporal para que pudesse ser feita a comparação entre o rastreamento fornecido pelo filtro de Kalman e o canal que deveria ser estimado. Para que pudesse ser feita uma comparação equivalente a feita

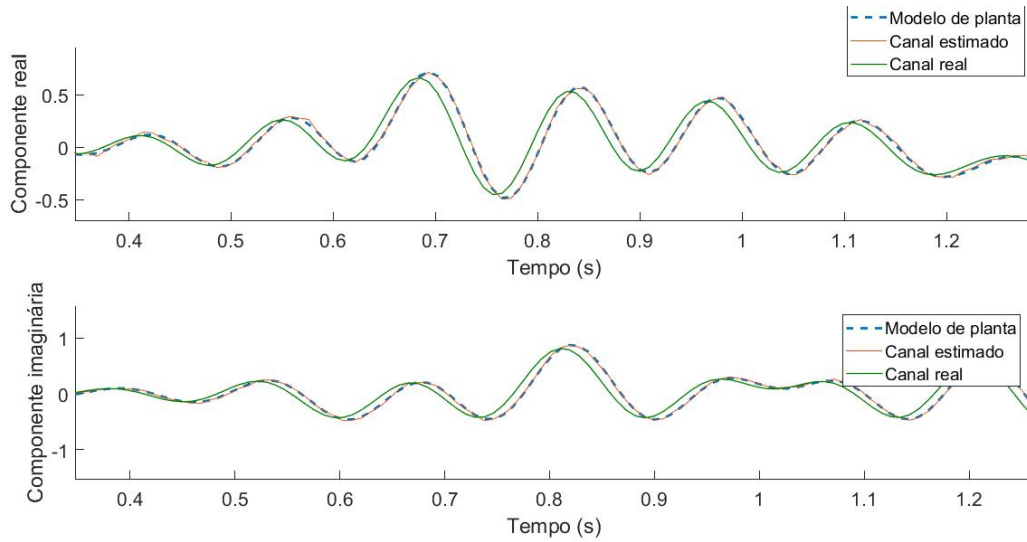


Figura 5.7: Corte feito no rastreamento do canal de comunicação com o uso filtro de Kalman.

no algoritmo LMS. Nota-se, neste caso, que a estimação com filtro de Kalman apresenta melhores resultados nos casos em que a SNR é baixa quando comparado ao caso LMS.

Apesar desse atraso temporal inerente a qualquer sistema causal, o filtro de Kalman apresentou melhores resultados mesmo em condições adversas como as em que a SNR é baixa. Um melhor desempenho é visto também quando há a presença de correlação entre sistema de comunicação como o mostrado em Fig. 5.11. Isso mostra que é possível um melhor rastreamento do canal mesmo em condições adversas, o que permite portanto uma ampliação dos benefícios da utilização de esquemas de comunicação MIMO.

## 5.6 Conclusões parciais

Nota-se que o filtro de Kalman apresentou melhores resultados em relação ao rastreo de canais de comunicação. Isso permite, como dito anteriormente, uma ampliação dos benefícios da utilização das esquemas MIMO massivo mesmo em situações em que haja grande correlação entre os canais ou em que a SNR é baixa.

O filtro de Kalman apresentou também um menor tempo de acomodação para que o rastreamento pudesse ser feito de maneira adequada. Além disso, nota-se que o problema da escolha de parâmetros para que o rastreamento funcione de maneira adequada são reduzidos. Tudo isso, no entanto, exige um modelamento adequado do processo no espaço de estados.

Em relação ao modelo usado para a descrição do canal no espaço de estados, nota-se que, quanto menor o coeficiente de desvanecimento temporal, melhor é a adequação desse ao canal. Isso porque um menor valor desse parâmetro significa que há um maior movimento relativo entre transmissor e receptor.

Foi possível ainda analisar o comportamento do filtro de Kalman mesmo em situações adversas

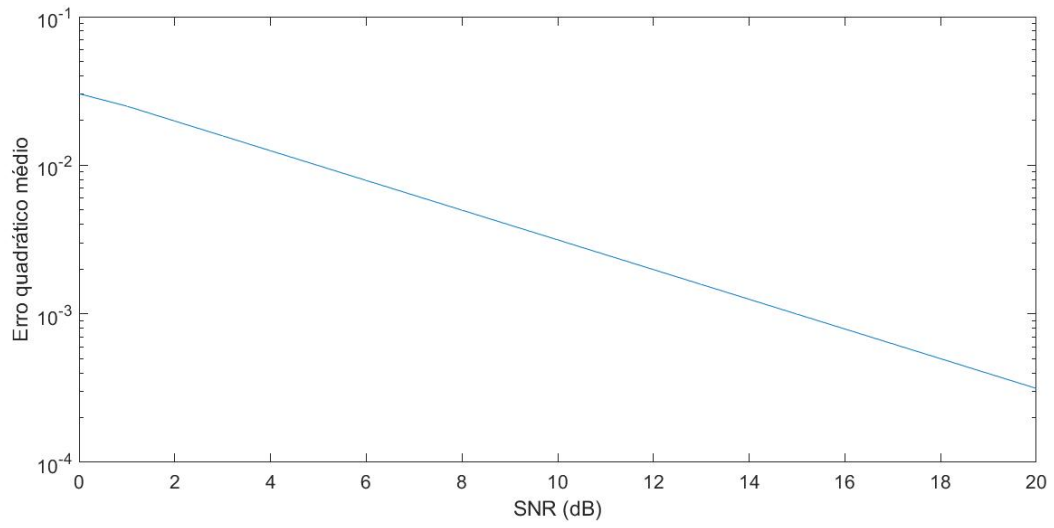


Figura 5.8: Erro quadrático médio entre saída do filtro de Kalman,  $y_e$ , e planta do processo,  $y$  pela SNR.

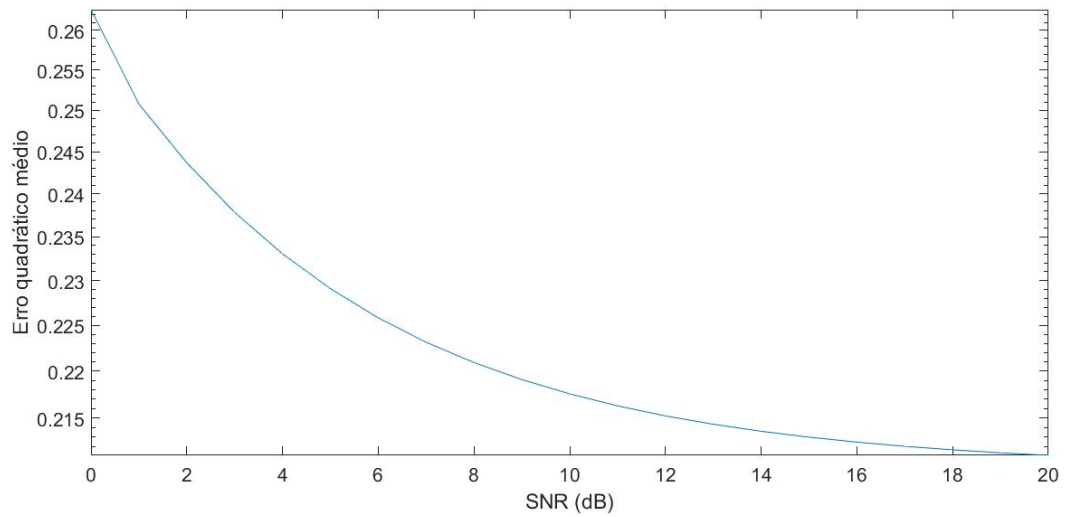


Figura 5.9: Erro quadrático médio entre saída do filtro de Kalman,  $y_e$  e canal real,  $u$ , pela SNR.

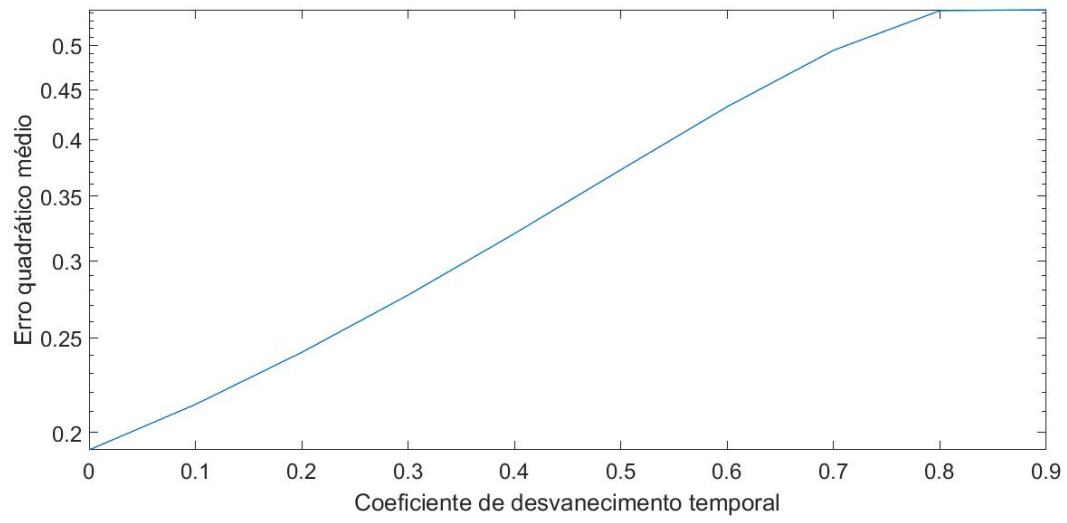


Figura 5.10: Erro quadrático médio pelo coeficiente de desvanecimento temporal.

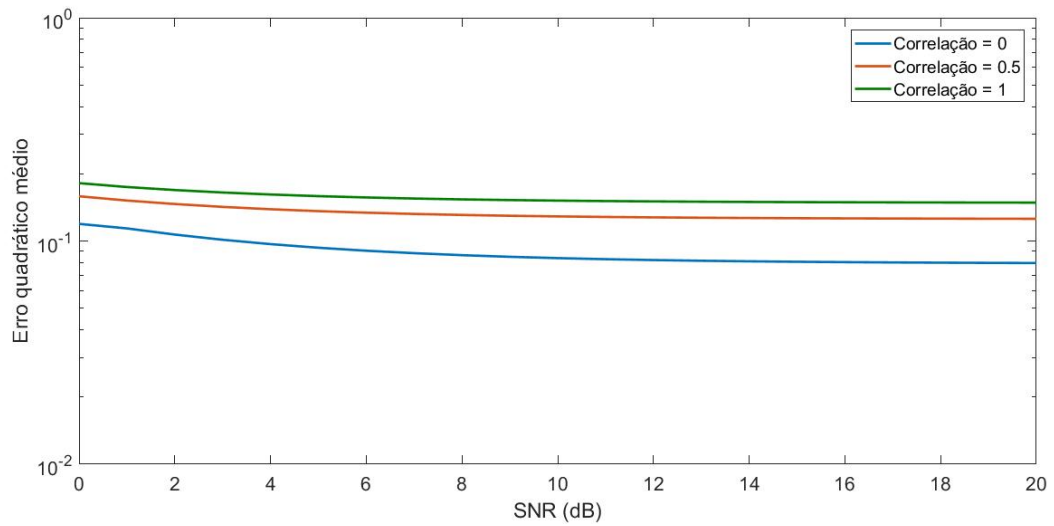


Figura 5.11: Erro quadrático médio entre saída do filtro de Kalman,  $y_e$ , e canal real,  $u$ , pela SNR.

tais como baixa relação sinal ruído e grande correlação entre os canais de comunicação. Esse é um fator presente em sistemas MIMO em que há uma grande quantidade de canais de comunicação. A geração de canais não correlacionados que foi o principal motivo para a mudança na abordagem da construção do canal de comunicação usado neste trabalho.

## Capítulo 6

# Conclusão

*Neste capítulo é apresentada a conclusão que resume os resultados obtidos no trabalho. Há ainda a presença de sugestões para propostas de trabalhos futuros.*

O desenvolvimento tecnológico exige cada vez mais um mundo conectado, em que o tráfego de dados demandado cresce de maneira muito rápida. E, como existem limitações ao aumento da largura de banda disponível e da potência de transmissão são necessárias novas tecnologias. A utilização de técnicas MIMO podem trazer essas mudanças necessárias. Nesse contexto, o correto rastreamento dos canais de comunicação permite um melhor conhecimento do canal de comunicação. Dessa forma, é possível que sistemas de comunicação sem fio apresentem melhor desempenho, que é cada vez mais demandado.

Para que se pudesse analisar o rastreamento dos algoritmos usados neste trabalho, LMS e filtro de Kalman, foi gerado um modelo de canal baseado no modelo de canal de Jakes. Este é um modelo desenvolvido anos atrás por pesquisadores de uma empresa destaque em inovação tecnológica e, apesar de sua representação bem simplificada, incorpora em sua descrição fenômenos como o desvio Doppler que é frequente quando se trata de comunicações móveis.

O algoritmo LMS que foi desenvolvido na década de 60, possui um algoritmo simples, não requer medidas da função da correlação do sistema nem a inversão da matriz de correlação. Além disso, esse algoritmo foi bastante utilizado principalmente na época em que foi desenvolvido. Seu funcionamento baseado na minimização instantânea do erro por acarretar, no entanto, problemas na convergência do problema quando a escolha do parâmetro de aprendizagem não é feita de maneira adequada.

Já o filtro de Kalman é um algoritmo computacional que existe a mais de 50 anos e ainda possui grande utilização atualmente. Exige um maior esforço computacional, porém sua convergência ocorre de maneira mais rápida e na maior parte das vezes. Com tudo isso, apresenta um desempenho superior mesmo em condições adversas como em sistemas cuja relação sinal ruído é baixa e quando há grande correlação entre os canais.

## 6.1 Comentários sobre os resultados obtidos

O objetivo principal do trabalho foi a análise do filtro de Kalman para o rastreamento de canais de comunicação frente a outro algoritmo de estimação. De maneira mais específica, essa análise foi feita em relação ao algoritmo LMS.

Notou-se que o filtro de Kalman possui rápida convergência enquanto o LMS possui um maior tempo de convergência. Dentro do algoritmo LMS esses tempo pode variar a depender da ordem do filtro e do parâmetro de aprendizado ajustados, apesar disso, o desempenho é inferior ao observado com o filtro de Kalman. Em sistemas que demandam uma rápida estimação do canal, esse tempo pode ser decisivo para haver um bom aproveitamento do canal de comunicação.

Além disso foi analisada a importância do correto modelamento no espaço do canal para que esse pudesse ser acompanhado de forma adequada. Isso porque a representação no espaço de estados é essencial para o correto funcionamento do filtro de Kalman. No modelo de canal percebe-se a importância do coeficiente de desvanecimento temporal para o rastreadimento. Além disso, quanto maior o máximo desvio Doppler, maior é portanto, o movimento relativo entre transmissor e receptor, o que gera um canal que varia rapidamente no tempo, o que resulta em um coeficiente de desvanecimento temporal do canal menor. Dessa forma, o valor atual estaria menos relacionado com o valor passado e mais relacionado com a inovação sofrida pelo canal.

## 6.2 Sugestões para trabalhos futuros

Como descrito diversas vezes durante o trabalho a correta estimação do canal é um ponto fundamental dentro de um sistema de comunicação sem fio para que essa se dê com maior qualidade, velocidade e robustez. E, um sistema de comunicação ideal é aquele em que os envolvidos transmissor e receptor tem informações sobre o canal de comunicação. Um sistema funcionando em malha aberta pode muitas vezes apresentar desempenho superior a um sistema que opera em malha fechada a depender do tempo ou da largura de banda usados para enviar informações sobre a matriz de ganhos do transmissor para o receptor. No entanto, dentro de qualquer uma das formas de operação o conhecimento do canal de comunicação é crucial para o melhor desempenho do canal.

O esquema que opera em malha fechada é mostrado em [20] e esse pode servir portanto como um também como um referencial para futuros trabalhos. Estes que são apresentados abaixo:

- Análise do rastreamento usando filtro de Kalman em outros modelos de canal de comunicação como o *Third Generation Partnership Project* (3GPP);
- Análise das não linearidades da cadeia de recepção no rastreamento obtido com o filtro de Kalman;
- A análise do rastreamento usando outras representações como uma que apresente variações de segunda ordem;

- Análise do impacto do erro de estimação na capacidade de um sistema;
- Análise do ganho de capacidade devido a um melhor rastreamento com um sistema em malha fechada mesmo que isso signifique diminuir a banda disponível ou o tempo de transmissão da informação;
- Dentro de um sistema que opera em malha fechada, a melhoria no rastreamento do canal de comunicação quando da utilização de livros códigos específicos em vez de livros código aleatório;



# REFERÊNCIAS BIBLIOGRÁFICAS

- [1] CHOI, J.; LOVE, D. J.; BIDIGARE, P. Downlink training techniques for fdd massive mimo systems: Open-loop and closed-loop training with memory. *IEEE Journal of Selected Topics in Signal Processing*, v. 8, n. 5, p. 802–814, Oct 2014. ISSN 1932-4553.
- [2] HAMPTON, J. *Introduction to MIMO Communications*. [S.l.]: Cambridge University Press, 2013. ISBN 9781107042834.
- [3] SANTOS, D. M. S. dos; ARAÚJO, D. C.; CAVALCANTE, C. C. Seleção de pilotos em sistemas de múltiplas antenas para maximização da snr recebida. *XXXIV Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT2016)*. Santarém-PA, 2016.
- [4] JAKES, W. *Microwave mobile communications*. [S.l.]: Wiley, 1974. (IEEE Press classic reissue). ISBN 9780471437208.
- [5] TRANTER, W. *Principles of Communication Systems Simulation with Wireless Applications*. [S.l.]: Prentice Hall Professional Technical Reference, 2004. (Prentice Hall communications engineering and emerging technologies series). ISBN 9780134947907.
- [6] SEMMLOW, J.; GRIFFEL, B. *Biosignal and Medical Image Processing, Third Edition*. [S.l.]: CRC Press, 2014. ISBN 9781466567375.
- [7] HAYKIN, S.; WIDROW, B. *Least-Mean-Square Adaptive Filters*. [S.l.]: Wiley, 2003. (Adaptive and Cognitive Dynamic Systems: Signal Processing, Learning, Communications and Control). ISBN 9780471215707.
- [8] FARAGHER, R. Understanding the basis of the kalman filter via a simple and intuitive derivation [lecture notes]. *IEEE Signal Processing Magazine*, v. 29, n. 5, p. 128–132, Sept 2012. ISSN 1053-5888.
- [9] RUSEK, F. et al. Scaling up mimo: Opportunities and challenges with very large arrays. *IEEE Signal Processing Magazine*, v. 30, n. 1, p. 40–60, Jan 2013. ISSN 1053-5888.
- [10] ITU. *Recommendation ITU-R M.1645: Vision, framework and overall objectives os the future development of IMT-2000 ans systems beyond IMG-2000*. [S.l.: s.n.].
- [11] LEITE, J. P. *Aplicação de Técnicas de Aprendizado por Reforço à Alocação de Recursos e ao Escalonamento de Usuários em Sistemas de telecomunicações*. Tese (Doutorado) — Universidade de Brasília, 2014.

- [12] LOIOLA, M. B. *Estimação de canais MIMO variantes no tempo usando filtros de Kalman*. Tese (Doutorado) — UNICAMP, 2009.
- [13] LARSSON, E. G. et al. Massive mimo for next generation wireless systems. *IEEE Communications Magazine*, v. 52, n. 2, p. 186–195, February 2014. ISSN 0163-6804.
- [14] MARZETTA, T. L. Noncooperative cellular wireless with unlimited numbers of base station antennas. *IEEE Transactions on Wireless Communications*, v. 9, n. 11, p. 3590–3600, November 2010. ISSN 1536-1276.
- [15] SILVA, M. F. *Técnicas de Estimação de Canal para Sistemas OFDM em Canais HF Baseadas no Arranjo COMB-TYPE de Subportadoras piloto*. Tese (Doutorado) — Universidade Federal do Rio de Janeiro, 2012.
- [16] GANS, M. J. A power-spectral theory of propagation in the mobile-radio environment. *IEEE Transactions on Vehicular Technology*, v. 21, n. 1, p. 27–38, Feb 1972. ISSN 0018-9545.
- [17] NISE, N. *Engenharia De Sistemas De Controle*. [S.l.]: LTC. ISBN 9788521621355.
- [18] ANDERSON, B.; MOORE, J. *Optimal Filtering*. [S.l.]: Dover Publications, 2012. (Dover Books on Electrical Engineering). ISBN 9780486136899.
- [19] SANTOS, D. M. S. dos. *Seleção de símbolos piloto em sistemas de comunicação sem fio*. Dissertação (Mestrado) — Universidade Federal do Ceará, 2016.
- [20] CHOI, J.; LOVE, D. J.; BIDIGARE, P. Downlink training techniques for fdd massive mimo systems: Open-loop and closed-loop training with memory. *IEEE Journal of Selected Topics in Signal Processing*, v. 8, n. 5, p. 802–814, Oct 2014. ISSN 1932-4553.

# ANEXOS

# I. DEMONSTRAÇÃO DAS EQUAÇÕES DO FILTRO DE KALMAN

## I.1 Estágio de predição

A relação entre a matriz de covariância e o erro é dada por Eq. (I.1), em que o erro é a diferença entre o valor real,  $x_{t|t}$ , e o valor estimado,  $\hat{x}_{t|t}$ . A matriz de covariância pode ser escrita então em Eq. (I.1). Objetivo com toda esta demonstração é encontrar uma forma de se expressar a matriz de covariância com base em valores passados no estágio de predição. Será mostrado também como é possível achar o ganho ideal mostrado em Eq. (5.31).

$$\mathbf{P}_{t|t} = E [(e_{t|t})(e_{t|t})^T] \quad (\text{I.1})$$

$$\mathbf{P}_{t|t} = E [(x_t - \hat{x}_{t|t})(x_t - \hat{x}_{t|t})^T] \quad (\text{I.2})$$

Pode-se substituir o valor visto em Eq. (5.28) para o valor estimado do vetor de estados,  $\hat{x}_t$ , em Eq. (I.2), com isso obtém-se Eq. (I.3)

$$\mathbf{P}_{t|t} = E [(\mathbf{x}_t - (\hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t(\mathbf{z}_t - \mathbf{H}_t\hat{\mathbf{x}}_{t|t-1})))(\mathbf{x}_t - (\hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t(\mathbf{z}_t - \mathbf{H}_t\hat{\mathbf{x}}_{t|t-1})))^T] \quad (\text{I.3})$$

Substituindo Eq. (5.30) em Eq. (I.3) e reorganizando os termos tem-se Eq. (I.4)

$$\mathbf{P}_{t|t} = E [(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t-1}) - \mathbf{K}_t(\mathbf{H}_t\mathbf{x}_t + \mathbf{V}_t - \mathbf{H}_t\hat{\mathbf{x}}_{t|t-1}))(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t-1}) - \mathbf{K}_t(\mathbf{H}_t\mathbf{x}_t + \mathbf{V}_t - \mathbf{H}_t\hat{\mathbf{x}}_{t|t-1}))^T] \quad (\text{I.4})$$

Reorganizando Eq. (I.4), tem-se Eq. (I.5)

$$\mathbf{P}_{t|t} = E [((\mathbf{x}_t - \hat{\mathbf{x}}_{t|t-1}) - \mathbf{K}_t(\mathbf{H}_t(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t-1}) + \mathbf{V}_t))((\mathbf{x}_t - \hat{\mathbf{x}}_{t|t-1}) - \mathbf{K}_t(\mathbf{H}_t(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t-1}) + \mathbf{V}_t))^T] \quad (\text{I.5})$$

Ao se fazer o produto distributivo em Eq. (I.5), observa-se que há grande semelhança entre Eq. (I.5) e Eq. (I.2). Na verdade, neste caso tem-se a matriz de covariância,  $\mathbf{P}_{t|t-1}$ , uma vez que o valor real é comparado com  $\hat{\mathbf{x}}_{t|t-1}$ . Além disso, o erro de medição,  $V_k$ , é não correlacionado com o erro de medição, pode-se escrever como em Eq. (I.6), em que  $R_t$  é definido como em Eq. (5.4) como sendo igual a  $R_t = E[v_tv_t^T]$ .

$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{K}_t\mathbf{H}_t\mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\mathbf{H}_t^T\mathbf{K}_t^T + \mathbf{K}_t(\mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^T + \mathbf{R}_t)\mathbf{K}_t^T \quad (\text{I.6})$$

O funcionamento do filtro de Kalman se baseia na minimização do erro quadrático médio utilizando a formulação de mínimos quadrados. Neste caso, o erro é um parâmetro dentro da função de covariância  $\mathbf{P}_{t|t}$ . Por isso, busca-se derivar os termos ao longo da diagonal principal dessa matriz, ou seja, o traço para que se encontre os pontos críticos, dessa forma, é possível encontrar o valor do ganho que minimiza esse erro denominado ganho de Kalman. Para se fazer essa derivação são necessárias duas fórmulas de diferenciação de matrizes, mostradas em Eq. (I.7) e Eq. (I.8). Na diferenciação de  $\mathbf{P}_t$  em relação a  $\mathbf{K}_t$  é preciso também se notar que o traço de  $\mathbf{P}_{t|t-1}\mathbf{H}_t^T\mathbf{K}_t^T$  é igual ao traço da sua transposta  $\mathbf{K}_t\mathbf{H}_t\mathbf{P}_{t|t-1}$ .

$$\frac{d[\text{traço}(\mathbf{AB})]}{d\mathbf{A}} = \mathbf{B}^T, \mathbf{AB} \text{ deve ser uma matriz quadrada} \quad (\text{I.7})$$

$$\frac{d[\text{traço}(\mathbf{ACA}^T)]}{d\mathbf{A}} = 2\mathbf{AC}, \mathbf{C} \text{ deve ser uma matriz simétrica} \quad (\text{I.8})$$

$$\frac{d\text{traço}(\mathbf{P}_t)}{d\mathbf{K}_t} = -2(\mathbf{H}_t\mathbf{P}_{t|t-1})^T + 2\mathbf{K}_t(\mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^T + \mathbf{R}_t) = 0 \quad (\text{I.9})$$

Assim, chega-se ao ganho de Kalman, mostrado em Eq. (I.10) e Eq. (5.31), responsável pela minimização do erro quadrático médio.

$$\mathbf{k}_t = \mathbf{P}_{t|t-1}\mathbf{H}_t^T (\mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^T + \mathbf{R}_t)^{-1} \quad (\text{I.10})$$

Substituindo o ganho de Kalman em Eq. (I.6) chega-se em Eq. (I.11), que é mostrado também em Eq. (5.29).

$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{K}_t\mathbf{H}_t\mathbf{P}_{t|t-1} \quad (\text{I.11})$$

## II. CONVERGÊNCIA DO ALGORITMO LMS

### II.1 Convergência na média

Considera-se que o vetor  $\mathbf{w}$  é não correlacionado com o vetor  $\mathbf{x}$  como mostrado em Eq. (II.1).

$$E[\hat{\mathbf{w}}(n)\mathbf{x}(n)] = 0 \quad (\text{II.1})$$

A partir de Eq. (4.25), pode-se aplicar o operador esperança o que gera Eq. (II.2) tendo como base a consideração feita em Eq. (II.1). Considera-se nesse caso que a matriz de correlação do vetor de entrada,  $\mathbf{R}_x$ , assume a forma definida em Eq. (I.3) e a matriz de correlação cruzada entre o vetor de entrada  $\mathbf{x}(n)$  e a resposta desejada  $\mathbf{d}(n)$ ,  $\mathbf{r}_{dx}$ , a apresentada em Eq. (I.4).

$$E[\hat{\mathbf{w}}(n+1)] = (\mathbf{I} - \eta\mathbf{R}_x)E[\hat{\mathbf{w}}(n)] + \eta\mathbf{r}_{dx} \quad (\text{II.2})$$

$$\mathbf{R}_x = E[\mathbf{x}(n)\mathbf{x}^T(n)] \quad (\text{II.3})$$

$$\mathbf{r}_{dx} = E[\mathbf{x}(n)\mathbf{d}(n)] \quad (\text{II.4})$$

Para demonstrar a convergência na média é necessário ainda fazer uso da transformação ortogonal de similaridade na matriz de correlação  $\mathbf{R}$ , como mostrado em Eq. (II.5), que pode ser reescrita como apresentado em Eq. (II.6). Nestas equações,  $\mathbf{A}$  é uma matriz diagonal cujos elementos são os autovalores da matriz de correlação  $\mathbf{R}_x$ ,  $\mathbf{Q}$  é uma matriz ortogonal cujas colunas são associadas aos autovetores  $\mathbf{R}_x$ .

$$\mathbf{Q}^t\mathbf{R}_x\mathbf{Q} = \mathbf{A} \quad (\text{II.5})$$

$$\mathbf{R}_x = \mathbf{Q}\mathbf{A}\mathbf{Q}^T \quad (\text{II.6})$$

Mostra-se ainda uma importante propriedade das matrizes ortogonais em Eq. (II.7). Essa característica é a de para uma matriz ortogonal sua matriz inversa e sua transposta são idênticas. Essa propriedade pode ser reescrita como em Eq. (II.8).

$$\mathbf{Q}^{-1} = \mathbf{Q}^T \quad (\text{II.7})$$

$$\mathbf{Q}\mathbf{Q}^T = \mathbf{I} \quad (\text{II.8})$$

Pode-se ainda reescrever as equações de Wiener-Hopf mostradas anteriormente em Eq. (4.12) como mostrado em Eq. (II.9), em que  $\mathbf{w}_o$  é o vetor de pesos ótimo do filtro de Wiener.

$$\mathbf{R}_x \mathbf{w}_o = \mathbf{r}_{dx} \quad (\text{II.9})$$

Substituindo Eq. (II.9) em Eq. (II.2) obtém-se Eq. (II.10).

$$E[\hat{\mathbf{w}}(n+1)] = (\mathbf{I} - \eta \mathbf{R}_x) E[\hat{\mathbf{w}}(n)] + \eta \mathbf{R}_x \mathbf{w}_o \quad (\text{II.10})$$

Por sua vez, substituindo Eq. (II.6) e Eq. (II.8) em Eq. (II.10), chega-se a Eq. (II.11), que pode ser reorganizada para obter a forma apresentada em Eq. (II.12).

$$E[\hat{\mathbf{w}}(n+1)] = (\mathbf{Q}\mathbf{Q}^T - \eta \mathbf{Q}\mathbf{A}\mathbf{Q}^T) E[\hat{\mathbf{w}}(n)] + \eta \mathbf{Q}\mathbf{A}\mathbf{Q}^T \mathbf{w}_o \quad (\text{II.11})$$

$$\mathbf{Q}^T E[\hat{\mathbf{w}}(n+1)] = (\mathbf{I} - \eta \mathbf{A}) \mathbf{Q}^T E[\hat{\mathbf{w}}(n)] + \eta \mathbf{A} \mathbf{Q}^T \mathbf{w}_o \quad (\text{II.12})$$

Definindo-se o vetor  $\mathbf{v}(n)$  como uma versão transformada do desvio entre a média dos pesos estimados,  $E[\hat{\mathbf{w}}(n)]$ , e a solução ótima de Wiener,  $\mathbf{w}_o$  mostrada em Eq. (II.13). Essa equação que pode ser reescrita como em Eq. (II.14).

$$\mathbf{v}(n) = \mathbf{Q}^T (E[\hat{\mathbf{w}}(n)] - \mathbf{w}_o) \quad (\text{II.13})$$

$$E[\mathbf{w}(n)] = \mathbf{Q}\mathbf{v}(n) + \mathbf{w}_o \quad (\text{II.14})$$

A partir de Eq. (II.12) e Eq. (II.14), pode-se escrever Eq. (II.15), que representa um sistema de equações diferenciais desacoplado homogêneo de primeira ordem como mostrado em Eq. (II.16). Em que  $\lambda_k$  são os autovalores da matriz de correlação,  $\mathbf{R}_x$  e  $v_k(n)$  é o  $k$ -ésimo elemento do vetor  $\mathbf{v}(n)$ . A solução dessa equação é apresentado em Eq. (II.17), em que  $v_k(0)$  é o valor inicial de  $v_k(n)$ .

$$\mathbf{v}(n+1) = (\mathbf{I} - \eta \mathbf{A}) \mathbf{v}(n) \quad (\text{II.15})$$

$$\mathbf{v}_k(n+1) = (1 - \eta \lambda_k) v_k(n), \quad k = 1, 2, \dots, p \quad (\text{II.16})$$

Para que o algoritmo LMS seja convergente na média independentemente do valor de  $v_k(0)$ , a condição apresentada em Eq. (II.18) deve ser satisfeita. Dessa maneira, a medida em que  $n$  tende a infinito,  $v_k(n)$  tende a zero como mostrado em Eq. (II.19).

$$\mathbf{v}_k(n) = (1 - \eta \lambda_k)^n v_k(0), \quad k = 1, 2, \dots, p \quad (\text{II.17})$$

$$|1 - \eta\lambda_k| < 1, \quad k = 1, 2, \dots, p \quad (\text{II.18})$$

$$v_k(n) \rightarrow 0, \quad \text{quando } n \rightarrow \infty \quad (\text{II.19})$$

A limitação apresenta em Eq. (II.18) para a convergência do algoritmo LMS estabelece também limites para o parâmetro da taxa de aprendizagem,  $\eta$ , que é apresentada em Eq. (II.20), em que  $\lambda_{max}$  é o maior autovalor da matriz de autocorrelação  $\mathbf{R}_x$ .

$$0 < \eta < \frac{2}{\lambda_{max}} \quad (\text{II.20})$$

## II.2 Convergência na média quadrática

A convergência na média quadrática, segundo [7], possui uma demonstração mais complicada. Por isso, é apresentada o resultado da convergência média quadrática de maneira simplificada em Eq. (II.21), em que  $Tr[\mathbf{R}_x]$  é o traço da matriz de correlação,  $\mathbf{R}_x$ . E, como o traço de  $\mathbf{R}_x$  é sempre maior do que  $\lambda_{max}$ , caso o algoritmo seja convergente na média quadrática ele também é convergente na média apesar de o inverso não ser verdade.

$$0 < \eta < \frac{2}{Tr[\mathbf{R}_x]} \quad (\text{II.21})$$

$$Tr[\mathbf{R}_x] = \sum \lambda_k \geq \lambda_{max} \quad (\text{II.22})$$

E, como o traço de uma matriz quadrada é igual a soma dos elementos de sua diagonal, em que o  $k$ -ésimo valor da diagonal da matriz  $\mathbf{R}_x$  é igual ao valor médio quadrático da entrada  $x(n)$ . O traço da matriz pode ser visto como a soma da potência total na entrada medida sobre todos os  $p$  sensores do filtro espacial do problema descrito no capítulo 4, isso é representado em Eq. (4.22).

$$0 < \eta < \frac{2}{\text{Potência total na entrada}} \quad (\text{II.23})$$